(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
27 December 2001 (27.12.2001)

PCT

(10) International Publication Number
WO 01/98342 A1

(51) International Patent Classification[7]: C07K 14/47, C12N 5/10, 5/16, 15/12, 15/63, 15/64

(21) International Application Number: PCT/US01/19929

(22) International Filing Date: 22 June 2001 (22.06.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/213,161	22 June 2000 (22.06.2000)	US
60/213,156	22 June 2000 (22.06.2000)	US

(71) Applicants *(for all designated States except US)*: **SMITHKLINE BEECHAM CORPORATION** [US/US]; One Franklin Plaza, Philadelphia, PA 19103 (US). **SMITHKLINE BEECHAM P.L.C.** [GB/GB]; New Horizons Court, Great West Road, Brentford, Middlesex TW8 9EP (GB). **GLAXO GROUP LIMITED** [GB/GB]; Glaxo Wellcome House, Berkley Avenue, Greenford, Middlesex UB6 0NN (GB).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **AGARWAL, Pankaj** [IN/US]; 251 West DeKalb Pike, King of Prussia, PA 19406 (US). **COGSWELL, John, P.** [US/US]; Five Moore Drive, Research Triangle Park, NC 27709 (US). **KABNIC, Karen, S.** [US/US]; 4138 Presidnetial Drive, Lafayette Hill, PA 19444 (US). **LAI, Ying-Ta** [—/US]; 516 Spruce Avenue, Upper Darby, PA 19082 (US). **MARTENSEN, Shelby, A.** [US/US]; Five Moore Drive, Research Triangle Park, NC 27709 (US). **MURDOCK, Paul, R.** [GB/GB]; New Frontiers Science Park South, Third Avenue, Harlow, Essex CM19 5AW (GB). **SMITH, Randall, F.** [US/US]; 4138 Presidential Drive, Lafayette Hill, PA 19444 (US). **STRUM, Jay, C.** [US/US]; Five Moore Drive, Research Triangle Park, NC 27709 (US). **XIANG, Zhaoying** [CN/US]; 2413 Ridgeway, Fort Lee, NJ 07024 (US). **XIE, Qing** [CN/US]; 310 Sawmill Lane, Horsham, PA 19044 (US). **RIZNI, Safia, K.** [PK/US]; 4617 Pine Street, Philadelphia, PA 19143 (US).

(74) Agents: **VENETIANER, Stephen** et al.; SmithKline Beecham Corporation, Corporate Intellectual Property, UW2220, 709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406-0939 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report*

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

— *with sequence listing part of description published separately in electronic form and available upon request from the International Bureau*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/98342 A1

(54) Title: NOVEL COMPOUNDS

(57) Abstract: Polypeptides and polynucleotides of the genes set forth in Table I and methods for producing such polypeptides by recombinant techniques are disclosed. Also disclosed are methods for utilizing polypeptides and polynucleotides of the genes set forth in Table I in diagnostic assays.

Novel Compounds

Field of Invention

This invention relates to newly identified polypeptides and polynucleotides encoding such polypeptides, to their use in diagnosis and in identifying compounds that may be agonists, antagonists that are potentially useful in therapy, and to production of such polypeptides and polynucleotides. The polynucleotides and polypeptides of the present invention also relate to proteins with signal sequences which allow them to be secreted extracellularly or membrane-associated (hereinafter often referred collectively as secreted proteins or secreted polypeptides).

Background of the Invention

The drug discovery process is currently undergoing a fundamental revolution as it embraces "functional genomics", that is, high throughput genome- or gene-based biology. This approach as a means to identify genes and gene products as therapeutic targets is rapidly superseding earlier approaches based on "positional cloning". A phenotype, that is a biological function or genetic disease, would be identified and this would then be tracked back to the responsible gene, based on its genetic map position.

Functional genomics relies heavily on high-throughput DNA sequencing technologies and the various tools of bioinformatics to identify gene sequences of potential interest from the many molecular biology databases now available. There is a continuing need to identify and characterise further genes and their related polypeptides/proteins, as targets for drug discovery.

Proteins and polypeptides that are naturally secreted into blood, lymph and other body fluids, or secreted into the cellular membrane are of primary interest for pharmaceutical research and development. The reason for this interest is the relative ease to target protein therapeutics into their place of action (body fluids or the cellular membrane). The natural pathway for protein secretion into extracellular space is the endoplasmic reticulum in eukaryotes and the inner membrane in prokaryotes (Palade, 1975, Science, 189, 347; Milstein, Brownlee, Harrison, and Mathews, 1972, Nature New Biol., 239, 117; Blobel, and Dobberstein, 1975, J. Cell. Biol., 67, 835). On the other hand, there is no known natural pathway for exporting a protein from the exterior of the cells into the cytosol (with the exception of pinocytosis, a mechanism of snake venom toxin intrusion into cells). Therefore targeting protein therapeutics into cells poses extreme difficulties.

The secreted and membrane-associated proteins include but are not limited

to all peptide hormones and their receptors (including but not limited to insulin, growth hormones, chemokines, cytokines, neuropeptides, integrins, kallikreins, lamins, melanins, natriuretic hormones, neuropsin, neurotropins, pituitiary hormones, pleiotropins, prostaglandins, secretogranins, selectins, thromboglobulins, thymosins), the breast and colon cancer gene products, leptin, the obesity gene protein and its receptors, serum albumin, superoxide dismutase, spliceosome proteins, 7TM (transmembrane) proteins also called as G-protein coupled receptors, immunoglobulins, several families of serine proteinases (including but not limited to proteins of the blood coagulation cascade, digestive enzymes), deoxyribonuclease I, etc.

Therapeutics based on secreted or membrane-associated proteins approved by FDA or foreign agencies include but are not limited to insulin, glucagon, growth hormone, chorionic gonadotropin, follicle stimulating hormone, luteinizing hormone, calcitonin, adrenocorticotropic hormone (ACTH), vasopressin, interleukines, interferones, immunoglobulins, lactoferrin (diverse products marketed by several companies), tissue-type plasminogen activator (Alteplase by Genentech), hyaulorindase (Wydase by Wyeth-Ayerst), dornase alpha (Pulmozyme\ by Genentech), Chymodiactin (chymopapain by Knoll), alglucerase (Ceredase by Genzyme), streptokinase (Kabikinase by Pharmacia) (Streptase by Astra), etc. This indicates that secreted and membrane-associated proteins have an established, proven history as therapeutic targets. Clearly, there is a need for identification and characterization of further secreted and membrane-associated proteins which can play a role in preventing, ameliorating or correcting dysfunction or disease, including but not limited to diabetes, breast-, prostate-, colon cancer and other malignant tumors, hyper- and hypotension, obesity, bulimia, anorexia, growth abnormalities, asthma, manic depression, dementia, delirium, mental retardation, Huntington's disease, Tourette's syndrome, schizophrenia, growth, mental or sexual development disorders, and dysfunctions of the blood cascade system including those leading to stroke. The proteins of the present invention which include the signal sequences are also useful to further elucidate the mechanism of protein transport which at present is not entirely understood, and thus can be used as research tools.

Summary of the Invention

The present invention relates to particular polypeptides and polynucleotides of the genes set forth in Table I, including recombinant materials and methods for their production.

Such polypeptides and polynucleotides are of interest in relation to methods of treatment of certain diseases, including, but not limited to, the diseases set forth in Tables III and V, hereinafter referred to as "diseases of the invention". In a further aspect, the invention relates to methods for identifying agonists and antagonists (*e.g.*, inhibitors) using the materials provided by the invention, and treating conditions associated with imbalance of polypeptides and/or polynucleotides of the genes set forth in Table I with the identified compounds. In still a further aspect, the invention relates to diagnostic assays for detecting diseases associated with inappropriate activity or levels the genes set forth in Table I. Another aspect of the invention concerns a polynucleotide comprising any of the nucleotide sequences set forth in the Sequence Listing and a polypeptide comprising a polypeptide encoded by the nucleotide sequence. In another aspect, the invention relates to a polypeptide comprising any of the polypeptide sequences set forth in the Sequence Listing and recombinant materials and methods for their production. Another aspect of the invention relates to methods for using such polypeptides and polynucleotides. Such uses include the treatment of diseases, abnormalities and disorders (hereinafter simply referred to as diseases) caused by abnormal expression, production, function and or metabolism of the genes of this invention, and such diseases are readily apparent by those skilled in the art from the homology to other proteins disclosed for each attached sequence. In still another aspect, the invention relates to methods to identify agonists and antagonists using the materials provided by the invention, and treating conditions associated with the imbalance with the identified compounds. Yet another aspect of the invention relates to diagnostic assays for detecting diseases associated with inappropriate activity or levels of the secreted proteins of the present invention.

Description of the Invention

In a first aspect, the present invention relates to polypeptides the genes set forth in Table I. Such polypeptides include:

(a) an isolated polypeptide encoded by a polynucleotide comprising a sequence set forth in the Sequence Listing, herein when referring to polynucleotides or polypeptides of the Sequence Listing, a reference is also made to the Sequence Listing referred to in the Sequence Listing;

(b) an isolated polypeptide comprising a polypeptide sequence having at least 95%, 96%, 97%, 98%, or 99% identity to a polypeptide sequence set forth in the Sequence Listing;

(c) an isolated polypeptide comprising a polypeptide sequence set forth in the Sequence Listing;

(d) an isolated polypeptide having at least 95%, 96%, 97%, 98%, or 99% identity to a polypeptide sequence set forth in the Sequence Listing;

(e) a polypeptide sequence set forth in the Sequence Listing; and

(f) an isolated polypeptide having or comprising a polypeptide sequence that has an Identity Index of 0.95, 0.96, 0.97, 0.98, or 0.99 compared to a polypeptide sequence set forth in the Sequence Listing;

(g) fragments and variants of such polypeptides in (a) to (f).

Polypeptides of the present invention are believed to be members of the gene families set forth in Table II . They are therefore of therapeutic and diagnostic interest for the reasons set forth in Tables III and V. The biological properties of the polypeptides and polynucleotides of the genes set forth in Table I are hereinafter referred to as "the biological activity" of polypeptides and polynucleotides of the genes set forth in Table I. Preferably, a polypeptide of the present invention exhibits at least one biological activity of the genes set forth in Table I.

Polypeptides of the present invention also include variants of the aforementioned polypeptides, including all allelic forms and splice variants. Such polypeptides vary from the reference polypeptide by insertions, deletions, and substitutions that may be conservative or non-conservative, or any combination thereof. Particularly preferred variants are those in which several, for instance from 50 to 30, from 30 to 20, from 20 to 10, from 10 to 5, from 5 to 3, from 3 to 2, from 2 to 1 or 1 amino acids are inserted, substituted, or deleted, in any combination.

Preferred fragments of polypeptides of the present invention include an isolated polypeptide comprising an amino acid sequence having at least 30, 50 or 100 contiguous amino acids from an amino acid sequence set forth in the Sequence Listing, or an isolated polypeptide comprising an amino acid sequence having at least 30, 50 or 100 contiguous amino acids truncated or deleted from an amino acid sequence set forth in the Sequence Listing. Preferred fragments are biologically active fragments that mediate the biological activity of polypeptides and polynucleotides of the genes set forth in Table I, including those with a similar activity or an improved activity, or with a decreased undesirable activity. Also preferred are those fragments that are antigenic or immunogenic in an animal, especially in a human.

Fragments of a polypeptide of the invention may be employed for producing the corresponding full-length polypeptide by peptide synthesis; therefore, these variants may be employed as intermediates for producing the full-length polypeptides of the invention. A polypeptide of the present invention may be in the form of the "mature" protein or may be a

part of a larger protein such as a precursor or a fusion protein. It is often advantageous to include an additional amino acid sequence that contains secretory or leader sequences, pro-sequences, sequences that aid in purification, for instance multiple histidine residues, or an additional sequence for stability during recombinant production.

Polypeptides of the present invention can be prepared in any suitable manner, for instance by isolation form naturally occurring sources, from genetically engineered host cells comprising expression systems (*vide infra*) or by chemical synthesis, using for instance automated peptide synthesizers, or a combination of such methods. Means for preparing such polypeptides are well understood in the art.

In a further aspect, the present invention relates to polynucleotides of the genes set forth in Table I. Such polynucleotides include:

(a) an isolated polynucleotide comprising a polynucleotide sequence having at least 95%, 96%, 97%, 98%, or 99% identity to a polynucleotide sequence set forth in the Sequence Listing;

(b) an isolated polynucleotide comprising a polynucleotide set forth in the Sequence Listing;

(c) an isolated polynucleotide having at least 95%, 96%, 97%, 98%, or 99% identity to a polynucleotide set forth in the Sequence Listing;

(d) an isolated polynucleotide set forth in the Sequence Listing;

(e) an isolated polynucleotide comprising a polynucleotide sequence encoding a polypeptide sequence having at least 95%, 96%, 97%, 98%, or 99% identity to a polypeptide sequence set forth in the Sequence Listing;

(f) an isolated polynucleotide comprising a polynucleotide sequence encoding a polypeptide set forth in the Sequence Listing;

(g) an isolated polynucleotide having a polynucleotide sequence encoding a polypeptide sequence having at least 95%, 96%, 97%, 98%, or 99% identity to a polypeptide sequence set forth in the Sequence Listing;

(h) an isolated polynucleotide encoding a polypeptide set forth in the Sequence Listing;

(i) an isolated polynucleotide having or comprising a polynucleotide sequence that has an Identity Index of 0.95, 0.96, 0.97, 0.98, or 0.99 compared to a polynucleotide sequence set forth in the Sequence Listing;

(j) an isolated polynucleotide having or comprising a polynucleotide sequence encoding a polypeptide sequence that has an Identity Index of 0.95, 0.96, 0.97, 0.98, or 0.99 compared to a polypeptide sequence set forth in the Sequence Listing; and

polynucleotides that are fragments and variants of the above mentioned polynucleotides or that are complementary to above mentioned polynucleotides, over the entire length thereof.

Preferred fragments of polynucleotides of the present invention include an isolated polynucleotide comprising an nucleotide sequence having at least 15, 30, 50 or 100 contiguous nucleotides from a sequence set forth in the Sequence Listing, or an isolated polynucleotide comprising a sequence having at least 30, 50 or 100 contiguous nucleotides truncated or deleted from a sequence set forth in the Sequence Listing.

Preferred variants of polynucleotides of the present invention include splice variants, allelic variants, and polymorphisms, including polynucleotides having one or more single nucleotide polymorphisms (SNPs).

Polynucleotides of the present invention also include polynucleotides encoding polypeptide variants that comprise an amino acid sequence set forth in the Sequence Listing and in which several, for instance from 50 to 30, from 30 to 20, from 20 to 10, from 10 to 5, from 5 to 3, from 3 to 2, from 2 to 1 or 1 amino acid residues are substituted, deleted or added, in any combination.

In a further aspect, the present invention provides polynucleotides that are RNA transcripts of the DNA sequences of the present invention. Accordingly, there is provided an RNA polynucleotide that:.

(a) comprises an RNA transcript of the DNA sequence encoding a polypeptide set forth in the Sequence Listing;

(b) is a RNA transcript of a DNA sequence encoding a polypeptide set forth in the Sequence Listing;

(c) comprises an RNA transcript of a DNA sequence set forth in the Sequence Listing; or

(d) is a RNA transcript of a DNA sequence set forth in the Sequence Listing; and RNA polynucleotides that are complementary thereto.

The polynucleotide sequences set forth in the Sequence Listing show homology with the polynucleotide sequences set forth in Table II. A polynucleotide sequence set forth in the Sequence Listing is a cDNA sequence that encodes a polypeptide set forth in the Sequence Listing. A polynucleotide sequence encoding a polypeptide set forth in the Sequence Listing may be identical to a polypeptide encoding a sequence set forth in the Sequence Listing or it may be a sequence other than a sequence set forth in the Sequence Listing, which, as a result of the redundancy (degeneracy) of the genetic code, also encodes a polypeptide set forth in the Sequence Listing. A polypeptide of a sequence set forth in the Sequence Listingis related to other proteins of the gene families set forth in Table II, having

homology and/or structural similarity with the polypeptides set forth in Table II. Preferred polypeptides and polynucleotides of the present invention are expected to have, *inter alia*, similar biological functions/properties to their homologous polypeptides and polynucleotides. Furthermore, preferred polypeptides and polynucleotides of the present invention have at least one activity of the genes set forth in Table I.

Polynucleotides of the present invention may be obtained using standard cloning and screening techniques from a cDNA library derived from mRNA from the tissues set forth in Table IV (see for instance, Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989)). Polynucleotides of the invention can also be obtained from natural sources such as genomic DNA libraries or can be synthesized using well known and commercially available techniques.

When polynucleotides of the present invention are used for the recombinant production of polypeptides of the present invention, the polynucleotide may include the coding sequence for the mature polypeptide, by itself, or the coding sequence for the mature polypeptide in reading frame with other coding sequences, such as those encoding a leader or secretory sequence, a pre-, or pro- or prepro- protein sequence, or other fusion peptide portions. For example, a marker sequence that facilitates purification of the fused polypeptide can be encoded. In certain preferred embodiments of this aspect of the invention, the marker sequence is a hexa-histidine peptide, as provided in the pQE vector (Qiagen, Inc.) and described in Gentz *et al.*, Proc Natl Acad Sci USA (1989) 86:821-824, or is an HA tag. A polynucleotide may also contain non-coding 5' and 3' sequences, such as transcribed, non-translated sequences, splicing and polyadenylation signals, ribosome binding sites and sequences that stabilize mRNA.

Polynucleotides that are identical, or have sufficient identity to a polynucleotide sequence set forth in the Sequence Listing, may be used as hybridization probes for cDNA and genomic DNA or as primers for a nucleic acid amplification reaction (for instance, PCR). Such probes and primers may be used to isolate full-length cDNAs and genomic clones encoding polypeptides of the present invention and to isolate cDNA and genomic clones of other genes (including genes encoding paralogs from human sources and orthologs and paralogs from other species) that have a high sequence similarity to sequences set forth in the Sequence Listing, typically at least 95% identity. Preferred probes and primers will generally comprise at least 15 nucleotides, preferably, at least 30 nucleotides and may have at least 50, if not at least 100 nucleotides. Particularly preferred probes will have between

30 and 50 nucleotides. Particularly preferred primers will have between 20 and 25 nucleotides.

A polynucleotide encoding a polypeptide of the present invention, including homologs from other species, may be obtained by a process comprising the steps of screening a library under stringent hybridization conditions with a labeled probe having a sequence set forth in the Sequence Listing or a fragment thereof, preferably of at least 15 nucleotides; and isolating full-length cDNA and genomic clones containing the polynucleotide sequence set forth in the Sequence Listing. Such hybridization techniques are well known to the skilled artisan. Preferred stringent hybridization conditions include overnight incubation at 42°C in a solution comprising: 50% formamide, 5xSSC (150mM NaCl, 15mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10 % dextran sulfate, and 20 microgram/ml denatured, sheared salmon sperm DNA; followed by washing the filters in 0.1x SSC at about 65°C. Thus the present invention also includes isolated polynucleotides, preferably with a nucleotide sequence of at least 100, obtained by screening a library under stringent hybridization conditions with a labeled probe having the sequence set forth in the Sequence Listing or a fragment thereof, preferably of at least 15 nucleotides.

The skilled artisan will appreciate that, in many cases, an isolated cDNA sequence will be incomplete, in that the region coding for the polypeptide does not extend all the way through to the 5' terminus. This is a consequence of reverse transcriptase, an enzyme with inherently low "processivity" (a measure of the ability of the enzyme to remain attached to the template during the polymerisation reaction), failing to complete a DNA copy of the mRNA template during first strand cDNA synthesis.

There are several methods available and well known to those skilled in the art to obtain full-length cDNAs, or extend short cDNAs, for example those based on the method of Rapid Amplification of cDNA ends (RACE) (see, for example, Frohman et al., Proc Nat Acad Sci USA 85, 8998-9002, 1988). Recent modifications of the technique, exemplified by the Marathon (trade mark) technology (Clontech Laboratories Inc.) for example, have significantly simplified the search for longer cDNAs. In the Marathon (trade mark) technology, cDNAs have been prepared from mRNA extracted from a chosen tissue and an 'adaptor' sequence ligated onto each end. Nucleic acid amplification (PCR) is then carried out to amplify the "missing" 5' end of the cDNA using a combination of gene specific and adaptor specific oligonucleotide primers. The PCR reaction is then repeated using 'nested' primers, that is, primers designed to anneal within the amplified product (typically an adapter specific primer that anneals further 3' in the adaptor sequence and a gene specific

primer that anneals further 5' in the known gene sequence). The products of this reaction can then be analyzed by DNA sequencing and a full-length cDNA constructed either by joining the product directly to the existing cDNA to give a complete sequence, or carrying out a separate full-length PCR using the new sequence information for the design of the 5' primer.

Recombinant polypeptides of the present invention may be prepared by processes well known in the art from genetically engineered host cells comprising expression systems. Accordingly, in a further aspect, the present invention relates to expression systems comprising a polynucleotide or polynucleotides of the present invention, to host cells which are genetically engineered with such expression systems and to the production of polypeptides of the invention by recombinant techniques. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention.

For recombinant production, host cells can be genetically engineered to incorporate expression systems or portions thereof for polynucleotides of the present invention. Polynucleotides may be introduced into host cells by methods described in many standard laboratory manuals, such as Davis et al., Basic Methods in Molecular Biology (1986) and Sambrook *et al.*(*ibid*). Preferred methods of introducing polynucleotides into host cells include, for instance, calcium phosphate transfection, DEAE-dextran mediated transfection, transvection, micro-injection, cationic lipid-mediated transfection, electroporation, transduction, scrape loading, ballistic introduction or infection.

Representative examples of appropriate hosts include bacterial cells, such as *Streptococci*, *Staphylococci*, *E. coli*, *Streptomyces* and *Bacillus subtilis* cells; fungal cells, such as yeast cells and *Aspergillus* cells; insect cells such as *Drosophila* S2 and *Spodoptera* Sf9 cells; animal cells such as CHO, COS, HeLa, C127, 3T3, BHK, HEK 293 and Bowes melanoma cells; and plant cells.

A great variety of expression systems can be used, for instance, chromosomal, episomal and virus-derived systems, *e.g.*, vectors derived from bacterial plasmids, from bacteriophage, from transposons, from yeast episomes, from insertion elements, from yeast chromosomal elements, from viruses such as baculoviruses, papova viruses, such as SV40, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and retroviruses, and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids. The expression systems may contain control regions that regulate as well as engender expression. Generally, any system or vector that is able to maintain, propagate or express a polynucleotide to produce a

polypeptide in a host may be used. The appropriate polynucleotide sequence may be inserted into an expression system by any of a variety of well-known and routine techniques, such as, for example, those set forth in Sambrook *et al.*, (*ibid*). Appropriate secretion signals may be incorporated into the desired polypeptide to allow secretion of the translated protein into the lumen of the endoplasmic reticulum, the periplasmic space or the extracellular environment. These signals may be endogenous to the polypeptide or they may be heterologous signals.

If a polypeptide of the present invention is to be expressed for use in screening assays, it is generally preferred that the polypeptide be produced at the surface of the cell. In this event, the cells may be harvested prior to use in the screening assay. If the polypeptide is secreted into the medium, the medium can be recovered in order to recover and purify the polypeptide. If produced intracellularly, the cells must first be lysed before the polypeptide is recovered.

Polypeptides of the present invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography is employed for purification. Well known techniques for refolding proteins may be employed to regenerate active conformation when the polypeptide is denatured during intracellular synthesis, isolation and/or purification.

Polynucleotides of the present invention may be used as diagnostic reagents, through detecting mutations in the associated gene. Detection of a mutated form of a gene is characterized by the polynucleotides set forth in the Sequence Listing in the cDNA or genomic sequence and which is associated with a dysfunction. Will provide a diagnostic tool that can add to, or define, a diagnosis of a disease, or susceptibility to a disease, which results from under-expression, over-expression or altered spatial or temporal expression of the gene. Individuals carrying mutations in the gene may be detected at the DNA level by a variety of techniques well known in the art.

Nucleic acids for diagnosis may be obtained from a subject's cells, such as from blood, urine, saliva, tissue biopsy or autopsy material. The genomic DNA may be used directly for detection or it may be amplified enzymatically by using PCR, preferably RT-PCR, or other amplification techniques prior to analysis. RNA or cDNA may also be used in similar fashion. Deletions and insertions can be detected by a change in size of the amplified product in comparison to the normal genotype. Point mutations can be identified

by hybridizing amplified DNA to labeled nucleotide sequences of the genes set forth in Table I. Perfectly matched sequences can be distinguished from mismatched duplexes by RNase digestion or by differences in melting temperatures. DNA sequence difference may also be detected by alterations in the electrophoretic mobility of DNA fragments in gels, with or without denaturing agents, or by direct DNA sequencing (see, for instance, Myers *et al.*, Science (1985) 230:1242). Sequence changes at specific locations may also be revealed by nuclease protection assays, such as RNase and S1 protection or the chemical cleavage method (see Cotton *et al.*, Proc Natl Acad Sci USA (1985) 85: 4397-4401).

An array of oligonucleotides probes comprising polynucleotide sequences or fragments thereof of the genes set forth in Table I can be constructed to conduct efficient screening of *e.g.*, genetic mutations. Such arrays are preferably high density arrays or grids. Array technology methods are well known and have general applicability and can be used to address a variety of questions in molecular genetics including gene expression, genetic linkage, and genetic variability, see, for example, M. Chee et al., Science, 274, 610-613 (1996) and other references cited therein.

Detection of abnormally decreased or increased levels of polypeptide or mRNA expression may also be used for diagnosing or determining susceptibility of a subject to a disease of the invention. Decreased or increased expression can be measured at the RNA level using any of the methods well known in the art for the quantitation of polynucleotides, such as, for example, nucleic acid amplification, for instance PCR, RT-PCR, RNase protection, Northern blotting and other hybridization methods. Assay techniques that can be used to determine levels of a protein, such as a polypeptide of the present invention, in a sample derived from a host are well-known to those of skill in the art. Such assay methods include radio-immunoassays, competitive-binding assays, Western Blot analysis and ELISA assays.

Thus in another aspect, the present invention relates to a diagnostic kit comprising:
(a) a polynucleotide of the present invention, preferably the nucleotide sequence set forth in the Sequence Listing , or a fragment or an RNA transcript thereof;
(b) a nucleotide sequence complementary to that of (a);
(c) a polypeptide of the present invention, preferably the polypeptide set forth in the Sequence Listing or a fragment thereof; or
(d) an antibody to a polypeptide of the present invention, preferably to the polypeptide set forth in the Sequence Listing .

It will be appreciated that in any such kit, (a), (b), (c) or (d) may comprise a substantial component. Such a kit will be of use in diagnosing a disease or susceptibility to a disease, particularly diseases of the invention, amongst others.

The polynucleotide sequences of the present invention are valuable for chromosome localisation studies. The sequences set forth in the Sequence Listing are specifically targeted to, and can hybridize with, a particular location on an individual human chromosome. The mapping of relevant sequences to chromosomes according to the present invention is an important first step in correlating those sequences with gene associated disease. Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. Such data are found in, for example, V. McKusick, Mendelian Inheritance in Man (available on-line through Johns Hopkins University Welch Medical Library). The relationship between genes and diseases that have been mapped to the same chromosomal region are then identified through linkage analysis (co-inheritance of physically adjacent genes). Precise human chromosomal localisations for a genomic sequence (gene fragment etc.) can be determined using Radiation Hybrid (RH) Mapping (Walter, M. Spillett, D., Thomas, P., Weissenbach, J., and Goodfellow, P., (1994) A method for constructing radiation hybrid maps of whole genomes, Nature Genetics 7, 22-28). A number of RH panels are available from Research Genetics (Huntsville, AL, USA) e.g. the GeneBridge4 RH panel (Hum Mol Genet 1996 Mar;5(3):339-46 A radiation hybrid map of the human genome. Gyapay G, Schmitt K, Fizames C, Jones H, Vega-Czarny N, Spillett D, Muselet D, Prud'Homme JF, Dib C, Auffray C, Morissette J, Weissenbach J, Goodfellow PN). To determine the chromosomal location of a gene using this panel, 93 PCRs are performed using primers designed from the gene of interest on RH DNAs. Each of these DNAs contains random human genomic fragments maintained in a hamster background (human / hamster hybrid cell lines). These PCRs result in 93 scores indicating the presence or absence of the PCR product of the gene of interest. These scores are compared with scores created using PCR products from genomic sequences of known location. This comparison is conducted at http://www.genome.wi.mit.edu/.

The polynucleotide sequences of the present invention are also valuable tools for tissue expression studies. Such studies allow the determination of expression patterns of polynucleotides of the present invention which may give an indication as to the expression patterns of the encoded polypeptides in tissues, by detecting the mRNAs that encode them. The techniques used are well known in the art and include in situ hydridization techniques to clones arrayed on a grid, such as cDNA microarray hybridization (Schena *et al*, Science, 270, 467-470, 1995 and Shalon *et al*, Genome Res, 6, 639-645, 1996) and nucleotide amplification techniques such as PCR. A preferred method uses the TAQMAN (Trade mark) technology available from Perkin Elmer. Results from these studies can provide an

indication of the normal function of the polypeptide in the organism. In addition, comparative studies of the normal expression pattern of mRNAs with that of mRNAs encoded by an alternative form of the same gene (for example, one having an alteration in polypeptide coding potential or a regulatory mutation) can provide valuable insights into the role of the polypeptides of the present invention, or that of inappropriate expression thereof in disease. Such inappropriate expression may be of a temporal, spatial or simply quantitative nature.

A further aspect of the present invention relates to antibodies. The polypeptides of the invention or their fragments, or cells expressing them, can be used as immunogens to produce antibodies that are immunospecific for polypeptides of the present invention. The term "immunospecific" means that the antibodies have substantially greater affinity for the polypeptides of the invention than their affinity for other related polypeptides in the prior art.

Antibodies generated against polypeptides of the present invention may be obtained by administering the polypeptides or epitope-bearing fragments, or cells to an animal, preferably a non-human animal, using routine protocols. For preparation of monoclonal antibodies, any technique which provides antibodies produced by continuous cell line cultures can be used. Examples include the hybridoma technique (Kohler, G. and Milstein, C., Nature (1975) 256:495-497), the trioma technique, the human B-cell hybridoma technique (Kozbor *et al.*, Immunology Today (1983) 4:72) and the EBV-hybridoma technique (Cole *et al.*, Monoclonal Antibodies and Cancer Therapy, 77-96, Alan R. Liss, Inc., 1985).

Techniques for the production of single chain antibodies, such as those described in U.S. Patent No. 4,946,778, can also be adapted to produce single chain antibodies to polypeptides of this invention. Also, transgenic mice, or other organisms, including other mammals, may be used to express humanized antibodies.

The above-described antibodies may be employed to isolate or to identify clones expressing the polypeptide or to purify the polypeptides by affinity chromatography. Antibodies against polypeptides of the present invention may also be employed to treat diseases of the invention, amongst others.

Polypeptides and polynucleotides of the present invention may also be used as vaccines. Accordingly, in a further aspect, the present invention relates to a method for inducing an immunological response in a mammal that comprises inoculating the mammal with a polypeptide of the present invention, adequate to produce antibody and/or T cell immune response, including, for example, cytokine-producing T cells or cytotoxic T cells,

to protect said animal from disease, whether that disease is already established within the individual or not. An immunological response in a mammal may also be induced by a method comprises delivering a polypeptide of the present invention *via* a vector directing expression of the polynucleotide and coding for the polypeptide *in vivo* in order to induce such an immunological response to produce antibody to protect said animal from diseases of the invention. One way of administering the vector is by accelerating it into the desired cells as a coating on particles or otherwise. Such nucleic acid vector may comprise DNA, RNA, a modified nucleic acid, or a DNA/RNA hybrid. For use a vaccine, a polypeptide or a nucleic acid vector will be normally provided as a vaccine formulation (composition). The formulation may further comprise a suitable carrier. Since a polypeptide may be broken down in the stomach, it is preferably administered parenterally (for instance, subcutaneous, intra-muscular, intravenous, or intra-dermal injection). Formulations suitable for parenteral administration include aqueous and non-aqueous sterile injection solutions that may contain anti-oxidants, buffers, bacteriostats and solutes that render the formulation instonic with the blood of the recipient; and aqueous and non-aqueous sterile suspensions that may include suspending agents or thickening agents. The formulations may be presented in unit-dose or multi-dose containers, for example, sealed ampoules and vials and may be stored in a freeze-dried condition requiring only the addition of the sterile liquid carrier immediately prior to use. The vaccine formulation may also include adjuvant systems for enhancing the immunogenicity of the formulation, such as oil-in water systems and other systems known in the art. The dosage will depend on the specific activity of the vaccine and can be readily determined by routine experimentation.

Polypeptides of the present invention have one or more biological functions that are of relevance in one or more disease states, in particular the diseases of the invention hereinbefore mentioned. It is therefore useful to identify compounds that stimulate or inhibit the function or level of the polypeptide. Accordingly, in a further aspect, the present invention provides for a method of screening compounds to identify those that stimulate or inhibit the function or level of the polypeptide. Such methods identify agonists or antagonists that may be employed for therapeutic and prophylactic purposes for such diseases of the invention as hereinbefore mentioned. Compounds may be identified from a variety of sources, for example, cells, cell-free preparations, chemical libraries, collections of chemical compounds, and natural product mixtures. Such agonists or antagonists so-identified may be natural or modified substrates, ligands, receptors, enzymes, etc., as the case may be, of the polypeptide; a structural or functional mimetic thereof (see Coligan *et al.*, Current Protocols in Immunology 1(2):Chapter 5 (1991)) or a small molecule. Such

small molecules preferably have a molecular weight below 2,000 daltons, more preferably between 300 and 1,000 daltons, and most preferably between 400 and 700 daltons. It is preferred that these small molecules are organic molecules.

The screening method may simply measure the binding of a candidate compound to the polypeptide, or to cells or membranes bearing the polypeptide, or a fusion protein thereof, by means of a label directly or indirectly associated with the candidate compound. Alternatively, the screening method may involve measuring or detecting (qualitatively or quantitatively) the competitive binding of a candidate compound to the polypeptide against a labeled competitor (*e.g.* agonist or antagonist). Further, these screening methods may test whether the candidate compound results in a signal generated by activation or inhibition of the polypeptide, using detection systems appropriate to the cells bearing the polypeptide. Inhibitors of activation are generally assayed in the presence of a known agonist and the effect on activation by the agonist by the presence of the candidate compound is observed. Further, the screening methods may simply comprise the steps of mixing a candidate compound with a solution containing a polypeptide of the present invention, to form a mixture, measuring an activity of the genes set forth in Table I in the mixture, and comparing activity of the mixture of the genes set forth in Table I to a control mixture which contains no candidate compound.

Polypeptides of the present invention may be employed in conventional low capacity screening methods and also in high-throughput screening (HTS) formats. Such HTS formats include not only the well-established use of 96- and, more recently, 384-well micotiter plates but also emerging methods such as the nanowell method described by Schullek et al, Anal Biochem., 246, 20-29, (1997).

Fusion proteins, such as those made from Fc portion and polypeptide of the genes set forth in Table I, as hereinbefore described, can also be used for high-throughput screening assays to identify antagonists for the polypeptide of the present invention (see D. Bennett *et al.*, J Mol Recognition, 8:52-58 (1995); and K. Johanson *et al.*, J Biol Chem, 270(16):9459-9471 (1995)).

The polynucleotides, polypeptides and antibodies to the polypeptide of the present invention may also be used to configure screening methods for detecting the effect of added compounds on the production of mRNA and polypeptide in cells. For example, an ELISA assay may be constructed for measuring secreted or cell associated levels of polypeptide using monoclonal and polyclonal antibodies by standard methods known in the art. This can be used to discover agents that may inhibit or enhance the production of polypeptide (also called antagonist or agonist, respectively) from suitably manipulated cells or tissues.

A polypeptide of the present invention may be used to identify membrane bound or soluble receptors, if any, through standard receptor binding techniques known in the art. These include, but are not limited to, ligand binding and crosslinking assays in which the polypeptide is labeled with a radioactive isotope (for instance, ^{125}I), chemically modified (for instance, biotinylated), or fused to a peptide sequence suitable for detection or purification, and incubated with a source of the putative receptor (cells, cell membranes, cell supernatants, tissue extracts, bodily fluids). Other methods include biophysical techniques such as surface plasmon resonance and spectroscopy. These screening methods may also be used to identify agonists and antagonists of the polypeptide that compete with the binding of the polypeptide to its receptors, if any. Standard methods for conducting such assays are well understood in the art.

Examples of antagonists of polypeptides of the present invention include antibodies or, in some cases, oligonucleotides or proteins that are closely related to the ligands, substrates, receptors, enzymes, etc., as the case may be, of the polypeptide, *e.g.*, a fragment of the ligands, substrates, receptors, enzymes, etc.; or a small molecule that bind to the polypeptide of the present invention but do not elicit a response, so that the activity of the polypeptide is prevented.

Screening methods may also involve the use of transgenic technology and the genes set forth in Table I. The art of constructing transgenic animals is well established. For example, the genes set forth in Table I may be introduced through microinjection into the male pronucleus of fertilized oocytes, retroviral transfer into pre- or post-implantation embryos, or injection of genetically modified, such as by electroporation, embryonic stem cells into host blastocysts. Particularly useful transgenic animals are so-called "knock-in" animals in which an animal gene is replaced by the human equivalent within the genome of that animal. Knock-in transgenic animals are useful in the drug discovery process, for target validation, where the compound is specific for the human target. Other useful transgenic animals are so-called "knock-out" animals in which the expression of the animal ortholog of a polypeptide of the present invention and encoded by an endogenous DNA sequence in a cell is partially or completely annulled. The gene knock-out may be targeted to specific cells or tissues, may occur only in certain cells or tissues as a consequence of the limitations of the technology, or may occur in all, or substantially all, cells in the animal. Transgenic animal technology also offers a whole animal expression-cloning system in which introduced genes are expressed to give large amounts of polypeptides of the present invention

Screening kits for use in the above described methods form a further aspect of the present invention. Such screening kits comprise:

(a) a polypeptide of the present invention;

(b) a recombinant cell expressing a polypeptide of the present invention;

(c) a cell membrane expressing a polypeptide of the present invention; or

(d) an antibody to a polypeptide of the present invention;

which polypeptide is preferably that set forth in the Sequence Listing.

It will be appreciated that in any such kit, (a), (b), (c) or (d) may comprise a substantial component.

Glossary

The following definitions are provided to facilitate understanding of certain terms used frequently hereinbefore.

"Antibodies" as used herein includes polyclonal and monoclonal antibodies, chimeric, single chain, and humanized antibodies, as well as Fab fragments, including the products of an

Fab or other immunoglobulin expression library.

"Isolated" means altered "by the hand of man" from its natural state, *i.e.*, if it occurs in nature, it has been changed or removed from its original environment, or both. For example, a polynucleotide or a polypeptide naturally present in a living organism is not "isolated," but the same polynucleotide or polypeptide separated from the coexisting materials of its natural state is "isolated", as the term is employed herein. Moreover, a polynucleotide or polypeptide that is introduced into an organism by transformation, genetic manipulation or by any other recombinant method is "isolated" even if it is still present in said organism, which organism may be living or non-living.

"Secreted protein activity or secreted polypeptide activity" or "biological activity of the secreted protein or secreted polypeptide" refers to the metabolic or physiologic function of said secreted protein including similar activities or improved activities or these activities with decreased undesirable side-effects. Also included are antigenic and immunogenic activities of said secreted protein.

"Secreted protein gene" refers to a polynucleotide comprising any of the attached nucleotide sequences or allelic variants thereof and/or their complements.

"Polynucleotide" generally refers to any polyribonucleotide (RNA) or polydeoxribonucleotide (DNA), which may be unmodified or modified RNA or DNA. "Polynucleotides" include, without limitation, single- and double-stranded DNA, DNA that

is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, "polynucleotide" refers to triple-stranded regions comprising RNA or DNA or both RNA and DNA. The term "polynucleotide" also includes DNAs or RNAs containing one or more modified bases and DNAs or RNAs with backbones modified for stability or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications may be made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically or metabolically modified forms of polynucleotides as typically found in nature, as well as the chemical forms of DNA and RNA characteristic of viruses and cells. "Polynucleotide" also embraces relatively short polynucleotides, often referred to as oligonucleotides.

"Polypeptide" refers to any polypeptide comprising two or more amino acids joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres. "Polypeptide" refers to both short chains, commonly referred to as peptides, oligopeptides or oligomers, and to longer chains, generally referred to as proteins. Polypeptides may contain amino acids other than the 20 gene-encoded amino acids. "Polypeptides" include amino acid sequences modified either by natural processes, such as post-translational processing, or by chemical modification techniques that are well known in the art. Such modifications are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature. Modifications may occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present to the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides may be branched as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched and branched cyclic polypeptides may result from post-translation natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, biotinylation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphotidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cystine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation,

myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination (see, for instance, Proteins - Structure and Molecular Properties, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York, 1993; Wold, F., Post-translational Protein Modifications: Perspectives and Prospects, 1-12, in Post-translational Covalent Modification of Proteins, B. C. Johnson, Ed., Academic Press, New York, 1983; Seifter *et al.*, "Analysis for protein modifications and nonprotein cofactors", Meth Enzymol, 182, 626-646, 1990, and Rattan *et al.*, "Protein Synthesis: Post-translational Modifications and Aging", Ann NY Acad Sci, 663, 48-62, 1992).

"Fragment" of a polypeptide sequence refers to a polypeptide sequence that is shorter than the reference sequence but that retains essentially the same biological function or activity as the reference polypeptide. "Fragment" of a polynucleotide sequence refers to a polynucleotide sequence that is shorter than the reference sequence set forth in the Sequence Listing.

"Variant" refers to a polynucleotide or polypeptide that differs from a reference polynucleotide or polypeptide, but retains the essential properties thereof. A typical variant of a polynucleotide differs in nucleotide sequence from the reference polynucleotide. Changes in the nucleotide sequence of the variant may or may not alter the amino acid sequence of a polypeptide encoded by the reference polynucleotide. Nucleotide changes may result in amino acid substitutions, additions, deletions, fusions and truncations in the polypeptide encoded by the reference sequence, as discussed below. A typical variant of a polypeptide differs in amino acid sequence from the reference polypeptide. Generally, alterations are limited so that the sequences of the reference polypeptide and the variant are closely similar overall and, in many regions, identical. A variant and reference polypeptide may differ in amino acid sequence by one or more substitutions, insertions, deletions in any combination. A substituted or inserted amino acid residue may or may not be one encoded by the genetic code. Typical conservative substitutions include Gly, Ala; Val, Ile, Leu; Asp, Glu; Asn, Gln; Ser, Thr; Lys, Arg; and Phe and Tyr. A variant of a polynucleotide or polypeptide may be naturally occurring such as an allele, or it may be a variant that is not known to occur naturally. Non-naturally occurring variants of polynucleotides and polypeptides may be made by mutagenesis techniques or by direct synthesis. Also included as variants are polypeptides having one or more post-translational modifications, for instance glycosylation, phosphorylation, methylation, ADP ribosylation and the like. Embodiments include methylation of the N-terminal amino acid, phosphorylations of serines and threonines and modification of C-terminal glycines.

"Allele" refers to one of two or more alternative forms of a gene occurring at a given locus in the genome.

"Polymorphism" refers to a variation in nucleotide sequence (and encoded polypeptide sequence, if relevant) at a given position in the genome within a population.

"Single Nucleotide Polymorphism" (SNP) refers to the occurrence of nucleotide variability at a single nucleotide position in the genome, within a population. An SNP may occur within a gene or within intergenic regions of the genome. SNPs can be assayed using Allele Specific Amplification (ASA). For the process at least 3 primers are required. A common primer is used in reverse complement to the polymorphism being assayed. This common primer can be between 50 and 1500 bps from the polymorphic base. The other two (or more) primers are identical to each other except that the final 3' base wobbles to match one of the two (or more) alleles that make up the polymorphism. Two (or more) PCR reactions are then conducted on sample DNA, each using the common primer and one of the Allele Specific Primers.

"Splice Variant" as used herein refers to cDNA molecules produced from RNA molecules initially transcribed from the same genomic DNA sequence but which have undergone alternative RNA splicing. Alternative RNA splicing occurs when a primary RNA transcript undergoes splicing, generally for the removal of introns, which results in the production of more than one mRNA molecule each of that may encode different amino acid sequences. The term splice variant also refers to the proteins encoded by the above cDNA molecules.

"Identity" reflects a relationship between two or more polypeptide sequences or two or more polynucleotide sequences, determined by comparing the sequences. In general, identity refers to an exact nucleotide to nucleotide or amino acid to amino acid correspondence of the two polynucleotide or two polypeptide sequences, respectively, over the length of the sequences being compared.

"% Identity" - For sequences where there is not an exact correspondence, a "% identity" may be determined. In general, the two sequences to be compared are aligned to give a maximum correlation between the sequences. This may include inserting "gaps" in either one or both sequences, to enhance the degree of alignment. A % identity may be determined over the whole length of each of the sequences being compared (so-called global alignment), that is particularly suitable for sequences of the same or very similar length, or over shorter, defined lengths (so-called local alignment), that is more suitable for sequences of unequal length.

"Similarity" is a further, more sophisticated measure of the relationship between two polypeptide sequences. In general, "similarity" means a comparison between the amino acids of two polypeptide chains, on a residue by residue basis, taking into account not only exact correspondences between a between pairs of residues, one from each of the sequences being compared (as for identity) but also, where there is not an exact correspondence, whether, on an evolutionary basis, one residue is a likely substitute for the other. This likelihood has an associated "score" from which the "% similarity" of the two sequences can then be determined.

Methods for comparing the identity and similarity of two or more sequences are well known in the art. Thus for instance, programs available in the Wisconsin Sequence Analysis Package, version 9.1 (Devereux J et al, Nucleic Acids Res, 12, 387-395, 1984, available from Genetics Computer Group, Madison, Wisconsin, USA), for example the programs BESTFIT and GAP, may be used to determine the % identity between two polynucleotides and the % identity and the % similarity between two polypeptide sequences. BESTFIT uses the "local homology" algorithm of Smith and Waterman (J Mol Biol, 147,195-197, 1981, Advances in Applied Mathematics, 2, 482-489, 1981) and finds the best single region of similarity between two sequences. BESTFIT is more suited to comparing two polynucleotide or two polypeptide sequences that are dissimilar in length, the program assuming that the shorter sequence represents a portion of the longer. In comparison, GAP aligns two sequences, finding a "maximum similarity", according to the algorithm of Neddleman and Wunsch (J Mol Biol, 48, 443-453, 1970). GAP is more suited to comparing sequences that are approximately the same length and an alignment is expected over the entire length. Preferably, the parameters "Gap Weight" and "Length Weight" used in each program are 50 and 3, for polynucleotide sequences and 12 and 4 for polypeptide sequences, respectively. Preferably, % identities and similarities are determined when the two sequences being compared are optimally aligned.

Other programs for determining identity and/or similarity between sequences are also known in the art, for instance the BLAST family of programs (Altschul S F et al, J Mol Biol, 215, 403-410, 1990, Altschul S F et al, Nucleic Acids Res., 25:389-3402, 1997, available from the National Center for Biotechnology Information (NCBI), Bethesda, Maryland, USA and accessible through the home page of the NCBI at www.ncbi.nlm.nih.gov) and FASTA (Pearson W R, Methods in Enzymology, 183, 63-99, 1990; Pearson W R and Lipman D J, Proc Nat Acad Sci USA, 85, 2444-2448,1988, available as part of the Wisconsin Sequence Analysis Package).

Preferably, the BLOSUM62 amino acid substitution matrix (Henikoff S and Henikoff J G, Proc. Nat. Acad Sci. USA, 89, 10915-10919, 1992) is used in polypeptide sequence comparisons including where nucleotide sequences are first translated into amino acid sequences before comparison.

Preferably, the program BESTFIT is used to determine the % identity of a query polynucleotide or a polypeptide sequence with respect to a reference polynucleotide or a polypeptide sequence, the query and the reference sequence being optimally aligned and the parameters of the program set at the default value, as hereinbefore described.

"Identity Index" is a measure of sequence relatedness which may be used to compare a candidate sequence (polynucleotide or polypeptide) and a reference sequence. Thus, for instance, a candidate polynucleotide sequence having, for example, an Identity Index of 0.95 compared to a reference polynucleotide sequence is identical to the reference sequence except that the candidate polynucleotide sequence may include on average up to five differences per each 100 nucleotides of the reference sequence. Such differences are selected from the group consisting of at least one nucleotide deletion, substitution, including transition and transversion, or insertion. These differences may occur at the 5' or 3' terminal positions of the reference polynucleotide sequence or anywhere between these terminal positions, interspersed either individually among the nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence. In other words, to obtain a polynucleotide sequence having an Identity Index of 0.95 compared to a reference polynucleotide sequence, an average of up to 5 in every 100 of the nucleotides of the in the reference sequence may be deleted, substituted or inserted, or any combination thereof, as hereinbefore described. The same applies *mutatis mutandis* for other values of the Identity Index, for instance 0.96, 0.97, 0.98 and 0.99.

Similarly, for a polypeptide, a candidate polypeptide sequence having, for example, an Identity Index of 0.95 compared to a reference polypeptide sequence is identical to the reference sequence except that the polypeptide sequence may include an average of up to five differences per each 100 amino acids of the reference sequence. Such differences are selected from the group consisting of at least one amino acid deletion, substitution, including conservative and non-conservative substitution, or insertion. These differences may occur at the amino- or carboxy-terminal positions of the reference polypeptide sequence or anywhere between these terminal positions, interspersed either individually among the amino acids in the reference sequence or in one or more contiguous groups within the reference sequence. In other words, to obtain a polypeptide sequence having an Identity Index of 0.95 compared to a reference polypeptide sequence, an average of up to 5

in every 100 of the amino acids in the reference sequence may be deleted, substituted or inserted, or any combination thereof, as hereinbefore described. The same applies *mutatis mutandis* for other values of the Identity Index, for instance 0.96, 0.97, 0.98 and 0.99.

The relationship between the number of nucleotide or amino acid differences and the Identity Index may be expressed in the following equation:

$$n_a \leq x_a - (x_a \bullet I),$$

in which:

n_a is the number of nucleotide or amino acid differences,

x_a is the total number of nucleotides or amino acids in a sequence set forth in the Sequence Listing,

I is the Identity Index,

• is the symbol for the multiplication operator, and

in which any non-integer product of x_a and I is rounded down to the nearest integer prior to subtracting it from x_a.

"Homolog" is a generic term used in the art to indicate a polynucleotide or polypeptide sequence possessing a high degree of sequence relatedness to a reference sequence. Such relatedness may be quantified by determining the degree of identity and/or similarity between the two sequences as hereinbefore defined. Falling within this generic term are the terms "ortholog", and "paralog". "Ortholog" refers to a polynucleotide or polypeptide that is the functional equivalent of the polynucleotide or polypeptide in another species. "Paralog" refers to a polynucleotideor polypeptide that within the same species which is functionally similar.

"Fusion protein" refers to a protein encoded by two, often unrelated, fused genes or fragments thereof. In one example, EP-A-0 464 533-A discloses fusion proteins comprising various portions of constant region of immunoglobulin molecules together with another human protein or part thereof. In many cases, employing an immunoglobulin Fc region as a part of a fusion protein is advantageous for use in therapy and diagnosis resulting in, for example, improved pharmacokinetic properties [see, *e.g.*, EP-A 0232 262]. On the other hand, for some uses it would be desirable to be able to delete the Fc part after the fusion protein has been expressed, detected and purified.

All publications and references, including but not limited to patents and patent applications, cited in this specification are herein incorporated by reference in their entirety as if each individual publication or reference were specifically and individually indicated to be incorporated by reference herein as being fully set forth. Any patent application to which

this application claims priority is also incorporated by reference herein in its entirety in the manner described above for publications and references.

Table I.

Gene Name	GSK Gene ID	Nucleic Acid SEQ ID NO's	Corresponding Protein SEQ ID NO's
sbg237163LIPASE	237163	SEQ ID NO:1	SEQ ID NO:23
sbg251170CEAa	251170	SEQ ID NO:2 SEQ ID NO:3	SEQ ID NO:24 SEQ ID NO:25
sbg389686WNT15a	389686	SEQ ID NO:4 SEQ ID NO:5	SEQ ID NO:26 SEQ ID NO:27
sbg236015LIPASE	236015	SEQ ID NO:6 SEQ ID NO:7	SEQ ID NO:28 SEQ ID NO:29
sbg417005LAMININ_ALPHA	417005	SEQ ID NO:8 SEQ ID NO:9	SEQ ID NO:30 SEQ ID NO:31
sbg425649KINASEa	425649	SEQ ID NO:10	SEQ ID NO:32
sbg419582PROTOCADHERIN	419582	SEQ ID NO:11 SEQ ID NO:12	SEQ ID NO:33 SEQ ID NO:34
sbg453915TECTORINa	453915	SEQ ID NO:13	SEQ ID NO:35
SBh385630.antiinflam	385630	SEQ ID NO:14 SEQ ID NO:15	SEQ ID NO:36 SEQ ID NO:37
sbg471005nAChR	471005	SEQ ID NO:16	SEQ ID NO:38
sbg442445PROa	442445	SEQ ID NO:17	SEQ ID NO:39
sbg456548CytoRa	456548	SEQ ID NO:18 SEQ ID NO:19	SEQ ID NO:40 SEQ ID NO:41
sbg456548CytoRa	456548b	SEQ ID NO:20	SEQ ID NO:42
sbg442358PROa	442358	SEQ ID NO:21 SEQ ID NO:22	SEQ ID NO:43 SEQ ID NO:44

Table II

Gene Name	Gene Family	Closest Polynuclotide by homology	Closest Polypeptide by homology	Cell Localization (by homology)
sbg237163 LIPASE	Pancreatic lipase	GB:AC011328 Direct submitted (06-OCT-1999) Genome Therapeutics Corporation, 100 Beaver Street, Waltham, MA 02453, USA	Mouse pancreatic lipase related protein 1, gi: 9256628 Remington,S.G., Lima,P.H. and Nelson,J.D. Invest. Ophthalmol. Vis. Sci. 40 (6), 1081-1090 (1999)	Secreted
sbg251170C EAa	Carcinoembryonic antigen	GB:AC020914 Submitted (12-JAN-2000) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA	Mouse putative protein, gi:12842545 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Genome Res. 10 (10), 1617-1630 (2000).	Secreted
sbg389686 WNT15a	WNT15	GB:AC015855 Directly submitted (17-NOV-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA.	Chicken WNT14 protein, gi:3915306 Bergstein I, Eisenberg LM, Bhalerao J, Jenkins NA, Copeland NG, Osborne MP, Bowcock AM, Brown AM; 1997; Genomics 46:450-8.	Secreted
sbg236015L IPASE	Lysosomal acid lipase	GB:AL358532 Directly submitted (15-DEC-2000) by Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK.	Rat lingual lipase, gi:126307 Docherty,A.J., Bodmer,M.W., Angal,S., Verger,R., Riviere,C., Lowe,P.A., Lyons,A., Emtage,J.S. and Harris,T.J. Nucleic Acids Res. 13 (6), 1891-1903 (1985)	Secreted
sbg417005L AMININ_ALPHA	Laminin alpha	GB:AL354836 Direct submitted (02-MAY-2000) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA	Human laminin alpha 5, gi:12274842 Submitted (14-FEB-2001) by Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK.	Secreted
sbg425649K INASEa	Casein kinase I-alpha	GB:AL356107 Submitted (16-MAY-2000) by Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK.	Human casein kinase I-alpha, gi:2134872 Fish,K.J., Cegielska,A., Getman,M.E., Landes,G.M. and Virshup,D.M. J. Biol. Chem. 270 (25), 14875-14883 (1995)	Cytosolic

sbg419582PROTOCADHERIN	Protocadherin	GB:AL355593 Direct submitted (17-MAY-2000) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK.	Human protocadherin 68 gi:11433373 Submitted (16-NOV-2000) by National Center for Biotechnology Information, NIH, Bethesda, MD 20894, USA	Secreted
sbg453915TECTORINa	Tectorin Beta	SC:AL157786 Submitted (04-MAY-2001) by Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK.	Mouse tectorin beta, gi:7363457 Legan,P.K., Rau,A., Keen,J.N. and Richardson,G.P. J. Biol. Chem. 272 (13), 8791-8801 (1997)	Secreted
SBh385630.antiinflam	Lipase	GB:AC015525 Submitted (16-NOV-1999) by Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA	Rabbit lacrimal lipase, gi:13560884 Submitted (20-FEB-2001) Ophthalmology, Regions Hospital, 640 Jackson Street, St. Paul, MN 55101, USA	Secreted

Table II (cont).

Gene Name	Gene Family	Closest Polynuclotide by homology	Closest Polypeptide by homology	Cell Localization (by homology)
sbg471005nAChR	Nicotinic acetylcholine receptor	GB:AC060812 Direct submitted (20-APR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA	Human cholinergic receptor, nicotinic, alpha polypeptide 10, gi:11138123 Lustig,L.R., Peng,H., Hiel,H., Yamamoto,T. and Fuchs,P.A. Genomics 73 (3), 272-283 (2001)	Membrane-bound
sbg442445PROa	Leucine rich repeat protein	GB:AC060234 Submitted (20-APR-2000) Genome Therapeutics Corporation, 100 Beaver Street, Waltham, MA 02453, USA	RIKEN cDNA mouse 4930442L21 gene Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Genome Res. 10 (10), 1617-1630 (2000)	Cytosolic
sbg456548CytoRa	Cytokine receptor	GB:AL158138 Submitted (20-JAN-2001) by Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK.	Human IL20 receptor, gi:7657691 Xie MH, Aggarwal S, Ho WH, Foster J, Zhang Z, Stinson J, Wood WI, Goddard AD and Gurney AL. J. Biol. Chem. 275 (40), 31335-31339 (2000)	Membrane-bound
sbg442358PROa	Leucine rich repeat protein	GB:AL139099 Submitted (23-MAY-2000) by Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE	Human EXMAD-9 geneseqp: AAB27231 Submitted by INCYTE GENOMICS INC Application and publication date: WO200068380-A2, 16-NOV-00	Membrane-bound

Table III

Gene Name	Uses	Associated Diseases
sbg237163 LIPASE	An embodiment of the invention is the use of sbg237163 LIPASE as replacement enzymes for patients with chronic pancreatitis. A close homologue of sbg237163 LIPASE is pancreatic lipase. Pancreatic lipase hydrolyzes dietary long chain triacylglycerol to free fatty acids and monoacylglycerols in the intestinal lumen (Lowe ME, Rosenblum JL, and Strauss AW; 1989; J Biol Chem 264:20042-8). Pancreatic steatorrhea and pancreatic diabetes are the dominant symptoms of patients in a certain stage of chronic pancreatitis. In this stage, the nutritional state is greatly disturbed and hypoglycemia and labile infection are involved. Pancreatic enzyme replacement therapy is the principal treatment method for pancreatic steatorrhea (Nakamura T, Takeuchi T, and Tando Y; 1998; Pancreas 16:329-36.	Cancer, infection, autoimmune disorder, hematopoietic disorder, wound healing disorders, inflammation.
sbg251170CEAa	An embodiment of the invention is the use of sbg251170CEAa as cell-surface molecules mediating cell-specific interactions in normal and neoplastic cells. A close homologue of sbg251170CEAa is carcinoembryonic antigen-related cell adhesion molecule 6. Carcinoembryonic antigen-related cell adhesion molecule 6 is claimed to function as a cell-surface molecules mediating cell-specific interactions in normal and neoplastic cells (1. Barnett T, Goebel SJ, Nothdurft MA, Elting JJ, Carcinoembryonic antigen family: characterization of cDNAs coding for NCA and CEA and suggestion of nonrandom sequence variation in their conserved loop-domains. Genomics 1988 Jul;3(1):59-66. 2. Inazawa J, Abe T, Inoue K, Misawa S, Oikawa S, Nakazato H, Yoshida MC. Regional assignment of nonspecific cross-reacting antigen (NCA) of the CEA gene family to chromosome 19 at band q13.2. Cytogenet Cell Genet 1989;52(1-2):28-31).	Cancer, autoimmune disorders, wound healing disorders, hematopoietic disorders and infection
sbg389686 WNT15a	An embodiment of the invention is the use of sbg389686WNT15a in regulation of cell growth and differentiation. Close homologues of sbg389686WNT15a are Wnt proteins. Wnt proteins are involved in critical developmental processes in both vertebrates and invertebrates and are implicated in regulation of cell growth and differentiation in certain adult mammalian tissues (Bergstein I, Eisenberg LM, Bhalerao J, Jenkins NA, Copeland NG, Osborne MP, Bowcock AM, Brown AM; 1997; Genomics 46:450-8). The Wnt gene family consists of at least 15 structurally related genes that encode secreted extracellular signaling factors. Wnt signaling is involved in many mammalian developmental processes, including cell proliferation, differentiation and epithelial-mesenchymal interactions, through which they contribute to the development of tissues and organs such as the limbs, the brain, the reproductive tract and the kidney. Evidence from tumor expression studies and transgenic animals experiments suggests that inappropriate activation of the Wnt signaling pathway is a major feature in human neoplasia and that oncogenic activation of this pathway can occur at many levels. Inappropriate expression of	Cancer, infection, autoimmune disorder, hematopoietic disorder, wound healing disorders, and inflammation

	the Wnt ligand and Wnt binding proteins have been found in a variety of human tumors (Smalley MJ, Dale TC;1999; Cancer Metastasis Rev 18:215-30).	
sbg236015LIPASE	An embodiment of the invention is the use of sbg236015LIPASE for treating lipase deficiency. A close homologue of sbg236015LIPASE is lysosomal acid lipase. The lysosomal acid lipase catalyzes the deacylation of triacylglyceryl and cholesteryl ester core lipids of endocytosed low density lipoproteins. This activity is deficient in patients with Wolman disease and cholesteryl ester storage disease, which are caused by a deficiency of lysosomal acid lipase activity, resulting in massive accumulation of cholesteryl ester and triglycerides (Anderson RA, Sando GN; 1991;J Biol Chem 266:22479-84).	Cancer, infection, autoimmune disorder, hematopoietic disorder, wound healing disorders, inflammation, Wolman disease, and cholesteryl ester storage disease
sbg417005LAMININ_ALPHA	An embodiment of the invention is the use of sbg417005LAMININ_ALPHA to promote myogenesis in skeletal muscle, outgrowth of neurites from central and peripheral neurons, and mesenchymal to epithelial transitions in kidney. A close homologue of sbg417005LAMININ_ALPHA is laminin. Laminins trimers, composed of alpha, beta, and gamma chains, are components of all basal laminae (BLs) throughout the bodies. In mammals they play at least three essential roles. First, they are major structural elements of BLs, forming one of two self-assembling networks to which other glycoproteins and proteoglycans of the BL attach. Second, they interact with cell surface components such as dystroglycan to attach cells to the extracellular matrix. Third, they are signaling molecules that interact with cellular receptors such as the integrins to convey important information to the cell interior. The alpha chains are ligands for most cellular laminin receptors. (Miner JH, Patton BL, Lentz SI, Gilbert DJ, Snider WD, Jenkins NA, Copeland NG, Sanes JR; 1997; J Cell Biol 137:685-701).	Cancer, infection, autoimmune disorder, hematopoietic disorder, wound healing disorders, inflammation, congenital muscular dystrophy, and junctional epidermolysis bullosa
sbg425649KINASEa	An embodiment of the invention is the use of sbg425649KINASEa in DNA replication and repair, membrane trafficking, neuroprotective, cytostatic, cardioactive, immunomodulatory, muscular, vulnerary, gastrointestinal, nephrotropic, anti-infective, gynaecological and antibacterial activities, and can be used in gene therapy. Close homologues of sbg425649KINASEa is mammalian casein kinases I (CKI) and human prostate cancer associated protein. CKI belongs to a family of serine/threonine protein kinases involved in diverse cellular processes including DNA replication and repair, membrane trafficking, circadian rhythms and Wnt signaling. Human prostate cancer associated proteins have neuroprotective, cytostatic, cardioactive, immunomodulatory, muscular, vulnerary, gastrointestinal, nephrotropic, anti-infective, gynaecological and antibacterial activities, and can be used in gene therapy.	Cancer, wound healing disorders, autoimmune disorders, hematopoietic disorders and infection

Table III (cont).

Gene Name	Uses	Associated Diseases
sbg419582PROTOCADHERIN	An embodiment of the invention is the use of sbg419582PROTOCADHERIN in functional systems of the nervous system, and may be involved in the formation of the neural network. A close homologue of sbg419582PROTOCADHERIN is protocadherin. The expression of protocadherin is developmentally regulated in a subset of the functional systems of the nervous system, and may be involved in the formation of the neural network by segregation of the brain nuclei and mediation of the axonal connections (Hirano S, Yan Q, Suzuki ST; 1999; J Neurosci 19:995-1005). The members of the cadherin superfamily are divided into two groups: classical cadherin type and protocadherin type. The current cadherins appear to have evolved from protocadherin (Suzuki ST; 1996; J Cell Sci 109:2609-11).	Cancer, infection, autoimmune disorder, hematopoietic disorder, wound healing disorders, inflammation, Parkinson's disease, Huntington's chorea, and multiple sclerosis
sbg453915TECTORINa	An embodiment of the invention is the use of sbg453915TECTORINa, a secreted protein, in cellular adhesion. A close homologue of sbg453915TECTORINa is mouse tectorin beta. The beta-tectorin is a protein of 36,074 Da that contains 4 consensus N glycosylation sites and a single zona pellucida domain. It is similar to components of the sperm-egg adhesion system, and, as such may have a similar functional role (Legan PK, Rau A, Keen JN, Richardson GP, The mouse tectorins. Modular matrix proteins of the inner ear homologous to components of the sperm-egg adhesion system. J Biol Chem 1997 Mar 28;272(13):8791-801).	Infection, cancer, wound healing disorders, hemotopoietic disorders and autoimmune disorders.
SBh385630.antiinflam	An embodiment of the invention is the use of SBh385630.antiinflam in gene therapy and are also suggested to have cytokine and cell proliferation/differentiation activity, immune stimulating (e.g.vaccines) or suppressing activity, haematopoiesis regulating activity, tissue growth activity, activin/inhibinactivity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, receptor/ligand activity,anti-inflammatory activity, cadherin/tumour invasion suppressor activity, and tumour inhibition activity. Lipases are also reported to be useful for gene therapy (WO9957132-A1;.Agostino, M.J., filed by GENETICS INST INC.). Close homologues of SBh385630.antiinflam include lipases.	Lematopoietic disorders, wound healing disorders, viral and bacterial infections, cancer, and autoimmune diseases
sbg471005nAChR	An embodiment of the invention is the use of sbg471005nAChR in physiological and behavioural processes of the brain. A close homologue of sbg471005nAChR is neuronal nicotinic acetylcholine receptors. Neuronal nicotinic acetylcholine receptors are a family of ion channels which are widely distributed in the human brain. There are many subtypes, and each has individual pharmacological and functional profiles. They mediate the effects of nicotine, and are involved in a number of physiological and behavioural processes. Additionally they may be implicated in a number of pathological conditions such	Cancer, infection, autoimmune disorder, hematopoietic disorder, wound healing disorders, inflammation, Alzheimer's disease, Parkinson's disease, and schizophrenia

	as Alzheimer's disease, Parkinson's disease and schizophrenia (Paterson D, Nordberg A; 2000; Prog Neurobiol 61:75-111).	
sbg442445PROa	An embodiment of the invention is the use of sbg442445PROa which may be involved in protein-protein interation and signal transduction in immune system. sbg442445PROa was expressed predominantly in lung and spleen/lymph. It encodes a protein with leucine rich repeats which may be involved in protein-protein interation and signal transduction in immune systems.	Inflammation, autoimmune disorders, asthma, allergies and sbg442445PROa-associated disorders
sbg456548CytoRa	The present gene has been cloned. Sybrman data showed its high expression levels in placenta and moderate levels in spleen and lymph. A close homologue of sbg456548CytoRa is another Class II cytokine receptor, ZCYTOR7. An embodiment of the invention is the use of sbg456548CytoRa, a decoy receptor, in the identification of other ligands, the promotion of anti-microbial activation of these cells, and/or potentiate the effectiveness of the natural ligand. Growth factors are known to promote the progression of cancer. A decoy receptor could interfere with that process. Proliferation, survival and differentiation can be transduced from activated cytokine receptors (Cell Signal. 1998. 10(9):619-628). Blocking these events could be crucial in modulating various diseases. The decoy receptor could potentially interfere with binding of these or other putative ligands, preventing downstream effects (Blood. 1999. 94(6):1943-1951). GM-CSF also has anti-apoptotic activity. A decoy receptor might then be able to block GM-CSF's anti-apoptotic actions when appropriate (Mol Biol Cell. 1999. 10(11):3959-3970). Roles for blocking the activity of the decoy receptor can be envisioned. GM-CSF promotes anti-microbial functions of mature neutrophils. Inhibiting the activity of an interfering decoy receptor could promote anti-microbial activation of these cells. Furthermore, rhGM-CSF is in wide clinical use to fight acute myeloid leukemia (Haematologica. 1991. 82(2): 239-245). Inhibition of a decoy receptor could potentiate the effectiveness of the natural ligand.	Chronic and acute inflammation, allergy, arthritis (including rheumatoid arthritis), septicemia, autoimmune diseases (e.g., inflammatory bowel disease, psoriasis), transplant rejection, graft vs. host disease, infection, stroke, ischemia, acute respiratory disease syndrome, asthma, restenosis, brain injury, AIDS, bone diseases, cancer, atheroschlerosis, Alzheimers disease, , hematopoietic disorder, and wound healing disorder
sbg442358PROa	An embodiment of the invention is the use of sbg442358PROa useful in the prevention and treatment of cancers, cell proliferation, cardiovascular, reproductive, immune, musculoskeletal, developmental and gastrointestinal disorders and inflammation. Close homologues of sbg442358PROa are human protein B27231 and Drosophila LRR47 that also contains leucine-rich repeats (LRRs) motifs. LRR has been found in a variety of extracellular, membrane and cytoplasmic proteins.and are believed to mediate specific protein-protein interactions and to function in cellular adhesion (Ntwasa,M., Buchanan,S.G. and Gay,N.J. Biochim. Biophys. Acta 1218 (2), 181-186 (1994)).	Cancer, autoimmune disorders, hemotopoietic disorders, wound healing disorders and infections

Table IV. Quantitative, Tissue-specific mRNA expression detected using SybrMan

Quantitative, tissue-specific, mRNA expression patterns of the genes were measured using SYBR-Green Quantitative PCR (Applied Biosystems, Foster City, CA; see Schmittgen T.D. et al., Analytical Biochemistry 285:194-204, 2000) and human cDNAs prepared from various human tissues. Gene-specific PCR primers were designed using the first nucleic acid sequence listed in the Sequence List for each gene. Results are presented as the number of copies of each specific gene's mRNA detected in 1ng mRNA pool from each tissue. Two replicate mRNA measurements were made from each tissue RNA.

Gene Name sbg237163LIPASE

Gene Name	Tissue-Specific mRNA Expression (copies per ng mRNA; avg. ± range for 2 data points per tissue)						
	Brain	**Heart**	**Lung**	**Liver**	**Kidney**	**Skeletal muscle**	**Intestine**
sbg237163LIPASE	5 ±0	8 ±2	7 ±2	-6 ±1	5 ±1	5 ±2	4 ±6

Gene Name sbg237163LIPASE cont.

Gene Name	Tissue-Specific mRNA Expression (copies per ng mRNA; avg. ± range for 2 data points per tissue)			
	Spleen/lymph	**Placenta**	**Testis**	
sbg237163LIPASE	3 ±2	1 ±1	47 ±1	

Gene Name sbg251170CEAa

Gene Name	Tissue-Specific mRNA Expression (copies per ng mRNA; avg. ± range for 2 data points per tissue)						
	Brain	**Heart**	**Lung**	**Liver**	**Kidney**	**Skeletal muscle**	**Intestine**
sbg250170CEAa	3 ±1	19 ±1	30 ±5	-5 ±3	3 ±1	5 ±5	21 ±2

Gene Name sbg251170CEAa cont.

Gene Name	Tissue-Specific mRNA Expression (copies per ng mRNA; avg. ± range for 2 data points per tissue)			
	Spleen/lymph	**Placenta**	**Testis**	
sbg237163LIPASE	33 ±4	22 ±3	14 ±0	

Table IV (cont).

In each gene's first subset table, two replicate measurements of gene of identification (GOI) mRNA were measured from various human tissues (column 2 and 3). The average GOI mRNA copies of the two replicates were made from each tissue RNA (column 4). The average amount of 18S rRNA from each tissue RNA was measured (column 5) and used for normalization. To make each tissue

with the same amount of 50 ng of 18S rRNA, the normalization factor (column 6) was calculated by dividing 50 ng with the amount of 18S rRNA measured from each tissue (column 5). The mRNA copies per 50 ng of total RNA were obtained by multipling each GOI normalization factor and average mRNA copies (column7).

Fold changes shown in each gene's second subset table were only calculated for disease tissues which have a normal counterpart. There are blanks in the fold change column for all samples that do not have counterparts. In addition, the fold change calculations are the fold change in the disease sample as compared to the normal sample. Accordingly, there will not be a fold change calculation next to any of the normal samples. For patient matched cancer pairs (colon, lung, and breast), each tumor is compared to its specific normal counterpart. When patient-matched normal/disease pairs do not exist, each disease sample was compared back to the average of all the normal samples of that same tissue type. For example, normal brain from the same patient that provided Alzheimer's brain is not applicable . Three normal brain samples and 4 Alzheimer's brain samples are used in the fold change. Three normal samples were averaged, and each of the Alzheimer's samples was compared back to that average.

Abbreviations

ALZ Alzheimer's Disease
CT CLONTECH (1020 East Meadow Circle Palo Alto, CA 94303-4230, USA)
KC Sample prepared by GSK investigator
COPD chronic obstructive pulmonary disease
endo endothelial
VEGF vascular endothelial growth factor
bFGF basic fibroblast growth factor
BM bone marrow
osteo osteoblast
OA osteoarthritis
RA rheumatoid arthritis
PBL peripheral blood lymphocytes
PBMNC peripheral blood mononuclear cells
HIV human immunodeficiency virus
HSV Herpes simplex virus
HPV human papilloma virus

Gene Name sbg389686WNT15a

Strong expression in Brain and dendritic cells. Brain expression may be from presence of glial cells. Expression in RA and OA synovium along with dendritic cells suggests a role for this protein in these diseases. Down regulation in ischemic and dilated heart indicates that replacement of protein could be therapeutic.

Sample sbg389686WNT15a	**Mean GOI copies (sample 1)**	**Mean GOI copies (sample 2)**	**Average GOI Copies**	**18S rRNA (ng)**	**50 ng/18S rRNA (ng)**	**copies of mRNA detected/ 50 ng total RNA**
Subcutaneous Adipocytes Zenbio	0.00	0.00	0.00	3.06	16.34	0.00
Subcutaneous Adipose Zenbio	0.00	1.71	0.86	0.96	52.36	44.76
Adrenal Gland Clontech	2.29	4.18	3.24	0.61	81.97	265.16
Whole Brain Clontech	698.52	625.01	661.77	7.24	6.91	4570.20
Fetal Brain Clontech	4.14	6.78	5.46	0.48	103.95	567.57

lung 24	-142.07
lung 28	-25.04
lung 23	-14.30
asthmatic lung	-2.44
asthmatic lung	1.41
asthmatic lung	3.06
asthmatic lung	2.31
endo VEGF	0.93
endo bFGF	5.16
heart T-1	1.90
heart T-14	1.20
heart T-3399	2.10
BM stim	-2.06
osteo undif	0.00
Cartilage (pool)	-1.97
PBL HIV IIIB	1.40
MRC5 HSV strain F	26.62

Gene Name sbg456548CytoRa

Strongly expressed in adenoid/tonsils and dendritic cells. Overexpressed in stimulated bone marrow. Taken together, these data suggest a role in immune function. Expression in GI tract suggests potential role in diseases of the GI system like IBD, Chron's, etc.

Sample sbg456548CytoRa	**Mean GOI copies (sample 1)**	**Mean GOI copies (sample 2)**	**Average GOI Copies**	**18S rRNA (ng)**	**50 ng/18S rRNA (ng)**	**copies of mRNA detected/ 50 ng total RNA**
Subcutaneous Adipocytes Zenbio	0.00	5.06	2.53	3.06	16.34	41.34
Subcutaneous Adipose Zenbio	0.00	0.00	0.00	0.96	52.36	0.00
Adrenal Gland Clontech	0.00	0.00	0.00	0.61	81.97	0.00
Whole Brain Clontech	0.00	0.00	0.00	7.24	6.91	0.00
Fetal Brain Clontech	0.00	0.00	0.00	0.48	103.95	0.00
Cerebellum Clontech	0.00	0.00	0.00	2.17	23.04	0.00
Cervix	0.00	7.86	3.93	2.42	20.66	81.20
Colon	9.12	37.61	23.37	2.71	18.45	431.09
Endometrium	0.00	0.00	0.00	0.73	68.21	0.00
Esophagus	0.00	0.00	0.00	1.37	36.50	0.00
Heart Clontech	0.00	0.00	0.00	1.32	37.88	0.00
Hypothalamus	0.00	0.00	0.00	0.32	155.28	0.00
Ileum	not done	39.63	39.63	2.58	19.38	768.02
Jejunum	9.16	33.67	21.42	6.60	7.58	162.23
Kidney	0.00	0.00	0.00	2.12	23.58	0.00
Liver	0.00	13.75	6.88	1.50	33.33	229.17
Fetal Liver Clontech	0.00	0.00	0.00	10.40	4.81	0.00
Lung	0.00	0.00	0.00	2.57	19.46	0.00

Mammary Gland Clontech	136.73	106.34	121.54	13.00	3.85	467.44
Myometrium	27.33	17.56	22.45	2.34	21.37	479.59
Omentum	0.00	12.61	6.31	3.94	12.69	80.01
Ovary	16.46	17.90	17.18	4.34	11.52	197.93
Pancreas	0.00	0.00	0.00	0.81	61.80	0.00
Head of Pancreas	0.00	0.00	0.00	1.57	31.85	0.00
Parotid Gland	21.25	23.72	22.49	5.48	9.12	205.16
Placenta Clontech	101.11	73.40	87.26	5.26	9.51	829.42
Prostate	8.55	0.00	4.28	3.00	16.67	71.25
Rectum	0.00	0.00	0.00	1.23	40.65	0.00
Salivary Gland Clontech	0.00	0.00	0.00	7.31	6.84	0.00
Skeletal Muscle Clontech	0.00	0.00	0.00	1.26	39.68	0.00
Skin	0.00	0.00	0.00	1.21	41.32	0.00
Small Intestine Clontech	0.00	0.00	0.00	0.98	51.07	0.00
Spleen	31.60	14.66	23.13	4.92	10.16	235.06
Stomach	0.00	7.01	3.51	2.73	18.32	64.19
Testis Clontech	0.00	0.00	0.00	0.57	87.87	0.00
Thymus Clontech	51.70	103.21	77.46	9.89	5.06	391.58
Thyroid	0.00	0.00	0.00	2.77	18.05	0.00
Trachea Clontech	0.00	0.00	0.00	9.71	5.15	0.00
Urinary Bladder	0.00	7.29	3.65	5.47	9.14	33.32
Uterus	5.98	21.02	13.50	5.34	9.36	126.40

Sample sbg456548CytoRa	**Reg number (GSK identifier)**	**Mean GOI copies**	**copies of mRNA detected/50 ng total RNA**	**Sample**	**Fold Change in Disease Population**
colon normal GW98-167	21941	54.19	108.38	colon normal	
colon tumor GW98-166	21940	242.87	485.74	colon tumor	4.481823215
colon normal GW98-178	22080	24.61	49.22	colon normal	
colon tumor GW98-177	22060	17.37	34.74	colon tumor	-1.416810593
colon normal GW98-561	23514	120.13	240.26	colon normal	
colon tumor GW98-560	23513	43.05	86.10	colon tumor	-2.79047619
colon normal GW98-894	24691	81.35	162.70	colon normal	
colon tumor GW98-893	24690	16.94	33.88	colon tumor	-4.802243211
lung normal GW98-3	20742	12.83	25.66	lung normal	
lung tumor GW98-2	20741	94.41	188.82	lung tumor	7.358534684
lung normal GW97-179	20677	519.7	1039.40	lung normal	
lung tumor GW97-178	20676	46.83	93.66	lung tumor	-11.09758702
lung normal GW98-165	21922	7.95	15.90	lung normal	
lung tumor GW98-164	21921	237.54	475.08	lung tumor	29.87924528
lung normal GW98-282	22584	251.04	502.08	lung normal	
lung tumor GW98-281	22583	28.16	56.32	lung tumor	-8.914772727
breast normal GW00-392	28750	138.99	138.99	breast normal	

breast tumor GW00-391	28746	147.66	295.32	breast tumor	2.124757177
breast normal GW00-413	28798	30.39	30.39	breast normal	
breast tumor GW00-412	28797	37.64	75.28	breast tumor	2.477130635
breast normal GW00-235:238	27592-95	218.09	218.09	breast normal	
breast tumor GW00-231:234	27588-91	14.68	14.68	breast tumor	-14.85626703
breast normal GW98-621	23656	1888.3	3776.60	breast normal	
breast tumor GW98-620	23655	877.2	1754.40	breast tumor	-2.152644779
brain normal BB99-542	25507	0	0.00	brain normal	
brain normal BB99-406	25509	0	0.00	brain normal	
brain normal BB99-904	25546	0	0.00	brain normal	
brain stage 5 ALZ BB99-874	25502	0	0.00	brain stage 5 ALZ	0
brain stage 5 ALZ BB99-887	25503	7.32	14.64	brain stage 5 ALZ	14.64
brain stage 5 ALZ BB99-862	25504	0	0.00	brain stage 5 ALZ	0
brain stage 5 ALZ BB99-927	25542	0	0.00	brain stage 5 ALZ	0
CT lung KC	normal	10.31	20.62	CT lung	
lung 26 KC	normal	49.79	49.79	lung 26	
lung 27 KC	normal	4.11	4.11	lung 27	
lung 24 KC	COPD	0.67	0.67	lung 24	-38.10074627
lung 28 KC	COPD	19.24	19.24	lung 28	-1.326793139
lung 23 KC	COPD	3.15	3.15	lung 23	-8.103968254
lung 25 KC	COPD	27.59	27.59	lung 25	
asthmatic lung ODO3112	29321	2.95	2.95	asthmatic lung	-8.653389831
asthmatic lung ODO3433	29323	9.86	19.72	asthmatic lung	-1.294497972
asthmatic lung ODO3397	29322	24.39	48.78	asthmatic lung	1.910880423
asthmatic lung ODO4928	29325	53.84	107.68	asthmatic lung	4.218196063
endo cells KC	control	0	0.00	endo cells	
endo VEGF KC		14.65	14.65	endo VEGF	14.65
endo bFGF KC		0	0.00	endo bFGF	0
heart Clontech	normal	0	0.00	heart	
heart (T-1) ischemic	29417	21.18	42.36	heart T-1	42.36
heart (T-14) non-obstructive DCM	29422	27.4	54.80	heart T-14	54.8
heart (T-3399) DCM	29426	93.27	186.54	heart T-3399	186.54
adenoid GW99-269	26162	579.69	1159.38	adenoid	
tonsil GW98-280	22582	3780.08	7560.16	tonsil	
T cells PC00314	28453	5.86	11.72	T cells	
PBMNC KC		0	0.00	PBMNC	
monocyte KC		0	0.00	monocyte	
B cells PC00665	28455	19.6	39.20	B cells	
dendritic cells 28441		580.67	1161.34	dendritic	

				cells	
neutrophils	28440	19.76	19.76	neutrophils	
eosinophils	28446	15.12	30.24	eosinophils	
BM unstim KC		0	0.00	BM unstim	
BM stim KC		296.72	296.72	BM stim	296.72
osteo dif KC		0	0.00	osteo dif	
osteo undif KC		0	0.00	osteo undif	0
chondrocytes		15.31	38.28	chondrocytes	
OA Synovium IP12/01	29462	39.57	39.57	OA Synovium	
OA Synovium NP10/01	29461	0	0.00	OA Synovium	
OA Synovium NP57/00	28464	70.08	140.16	OA Synovium	
RA Synovium NP03/01	28466	23.73	47.46	RA Synovium	
RA Synovium NP71/00	28467	24.13	48.26	RA Synovium	
RA Synovium NP45/00	28475	51.88	103.76	RA Synovium	
OA bone (biobank)	29217	0	0.00	OA bone (biobank)	
OA bone Sample 1	J. Emory	0	0.00	OA bone	
OA bone Sample 2	J. Emory	5.45	10.90	OA bone	
Cartilage (pool)	Normal	0	0.00	Cartilage (pool)	
Cartilage (pool)	OA	0	0.00	Cartilage (pool)	0
PBL unifected	28441	76.67	153.34	PBL unifected	
PBL HIV IIIB	28442	13.77	27.54	PBL HIV IIIB	-5.567901235
MRC5 uninfected (100%)	29158	0	0.00	MRC5 uninfected (100%)	
MRC5 HSV strain F	29178	0	0.00	MRC5 HSV strain F	0
W12 cells	29179	0	0.00	W12 cells	
Keratinocytes	29180	0	0.00	Keratinocytes	

Gene Name sbg456548CytoRa

Disease tissues	**Fold Change in Disease Population Relative to Normal**
colon tumor	4.48
colon tumor	-1.42
colon tumor	-2.79
colon tumor	-4.80
lung tumor	7.36

lung tumor	-11.10
lung tumor	29.88
lung tumor	-8.91
breast tumor	2.12
breast tumor	2.48
breast tumor	-14.86
breast tumor	-2.15
brain stage 5 ALZ	0.00
brain stage 5 ALZ	14.64
brain stage 5 ALZ	0.00
brain stage 5 ALZ	0.00
lung 24	-38.10
lung 28	-1.33
lung 23	-8.10
asthmatic lung	-8.65
asthmatic lung	-1.29
asthmatic lung	1.91
asthmatic lung	4.22
endo VEGF	14.65
endo bFGF	0.00
heart T-1	42.36
heart T-14	54.80
heart T-3399	186.54
BM stim	296.72
osteo undif	0.00
Cartilage (pool)	0.00
PBL HIV IIIB	-5.57
MRC5 HSV strain F	0.00

Gene Name sbg442358PROa

Expression in multiple immune cell types as well as stimulated bone marrow and thymus strongly suggests function in immune system. Overexpressed in breast tumors (1/4). Expression in RA and OA with corroborating expression in immune cells suggests role in these diseases. Overexpressed in heart disease suggesting role in CV diseases. Downregulated in HSV infected cells suggesting possible host cell factor.

Sample sbg442358PROa	**Mean GOI copies (sample 1)**	**Mean GOI copies (sample 2)**	**Average GOI Copies**	**18S rRNA (ng)**	**50 ng/18S rRNA (ng)**	**copies of mRNA detected/50 ng total RNA**
Subcutaneous Adipocytes Zenbio	1.86	1.71	1.79	3.06	16.34	29.17
Subcutaneous Adipose Zenbio	0.71	0.73	0.72	0.96	52.36	37.70

What is claimed is:

1. An isolated polypeptide selected from the group consisting of:
(a) an isolated polypeptide encoded by a polynucleotide comprising a sequence set forth in Table I;
(b) an isolated polypeptide comprising a polypeptide sequence set forth in Table I; and
(c) a polypeptide sequence of a gene set forth in Table I.

2. An isolated polynucleotide selected from the group consisting of:
(a) an isolated polynucleotide comprising a polynucleotide sequence set forth in Table I;
(b) an isolated polynucleotide of a gene set forth in Table I;
(c) an isolated polynucleotide comprising a polynucleotide sequence encoding a polypeptide set forth in Table I;
(d) an isolated polynucleotide encoding a polypeptide set forth in Table I;
(e) a polynucleotide which is an RNA equivalent of the polynucleotide of (a) to (d);
or a polynucleotide sequence complementary to said isolated polynucleotide.

3. An expression vector comprising a polynucleotide capable of producing a polypeptide of claim 1 when said expression vector is present in a compatible host cell.

4. A process for producing a recombinant host cell which comprises the step of introducing an expression vector comprising a polynucleotide capable of producing a polypeptide of claim 1 into a cell such that the host cell, under appropriate culture conditions, produces said polypeptide.

5. A recombinant host cell produced by the process of claim 4.

6. A membrane of a recombinant host cell of claim 5 expressing said polypeptide.

7. A process for producing a polypeptide which comprises culturing a host cell of claim 5 under conditions sufficient for the production of said polypeptide and recovering said polypeptide from the culture.

aaagccgacg cgcagcctcc aggttccgcc agcaccaacg tggtcctgcg ccacgatggc 420
gccgtgcgct gggacgcgcc ggccatcacg cgcagctcgt gccgcgtgga tgtagcagcc 480
ttcccgttcg acgcccagca ctgcggcctg acgttcggct cctggactca cggcgggcac 540
caactggatg tgcggccgcg cggcgctgca gccagcctgg cggacttcgt ggagaacgtg 600
gagtggcgcg tgctgggcat gccggcgcgg cggcgcgtgc tcacctacgg ctgctgctcc 660
gagccctacc ccgacgtcac cttcacgctg ctgctgcgcc gccgcgccgc cgcctacgtg 720
tgcaacctgc tgctgccctg cgtgctcatc tcgctgcttg cgccgctcgc cttccacctg 780
cctgccgact caggcgagaa ggtgtcgctg ggcgtcaccg tgctgctggc gctcaccgtc 840
ttccagttgc tgctggccga gagcatgcca ccggccgaga gcgtgccgct catcgggaag 900
tactacatgg ccactatgac catggtcaca ttctcaacag cactcaccat ccttatcatg 960
aacctgcatt actgtggtcc cagtgtccgc ccagtgccag cctgggctag ggccctcctg 1020
ctgggacacc tggcacgggg cctgtgcgtg cgggaaagag gggagccctg tgggcagtcc 1080
aggccacctg agttatctcc tagcccccag tcgcctgaag gaggggctgg ccccccagcg 1140
ggcccttgcc acgagccacg atgtctgtgc cgccaggaag ccctactgca ccacgtagcc 1200
accattgcca ataccttccg cagccaccga gctgcccagc gctgccatga ggactggaag 1260
cgcctggccc gtgtgatgga ccgcttcttc ctggccatct tcttctccat ggccctggtc 1320
atgagcctcc tggtgctggt gcaggccctg tga 1353

<210> 17
<211> 768
<212> DNA
<213> Homo sapiens

<400> 17
atggttaagg gtgagaaagg ccccaagggc aagaagatca ccctcaaggt ggccaggaat 60
tgcatcaaaa tcacttttga tgggaaaaag cgccttgact tgagcaagat gggaattacc 120
accttcccca agtgtattct gcgccttagt gacatggacg agctggacct tagccggaat 180
cttatcagga agatccctga ctccatctcc aagttccaga acctccggtg gctggacctg 240
cacagcaact acatagacaa gctgcctgag tccattggcc agatgaccag cctgctctac 300
ctcaacgtca gcaacaaccg gctgaccagc aacgggctgc ccgtggagct gaagcaactc 360
aagaacatcc gcgctgtgaa cctaggcttg aaccacctgg acagcgtgcc caccacactg 420
ggggccctga aggagctcca cgaggtaggg ctccatgaca acctactgaa caacatcccc 480
gtgagcatct ccaagctccc caagctgaaa aagctcaaca taaagcggaa ccccttcca 540
aagccaggtg agtcggaaat attcatagac tccatcagga ggctggagaa cttgtatgtt 600
gtggaggaga aggatctgtg tgcggcttgc ctgagaaaat gccaaaacgc ccgggacaac 660
ctgaatagaa tcaagaacat ggccacgacg acaccgagaa agaccatctt tcccaatctg 720
atctcaccca attccatggc caaggactcc tgggaagact ggaggtga 768

<210> 18
<211> 645
<212> DNA
<213> Homo sapiens

<400> 18
atgcaggcag gaactcagtc aacgcatgag tctctgaagc ctcagagggt acaatttcag 60
tcccgaaatt ttcacaacat tttgcaatgg cagcctggga gggcacttac tggcaacagc 120
agtgtctatt ttgtgcagta caaaatatat ggacagagac aatggaaaaa taaagaagac 180
tgttggggta ctcaagaact ctcttgtgac cttaccagtg aaacctcaga catacaggaa 240
ccttattacg ggagggtgag ggcggcctcg gctgggagct actcagaatg gagcatgacg 300
ccgcggttca ctccctggtg ggaaacaaaa atagatcctc cagtcatgaa tataacccaa 360
gtcaatggct ctttgttggt aattctccat gctccaaatt taccatatag ataccaaaag 420
gaaaaaaatg tatctataga agattactat gaactactat accgagtttt tataattaac 480
aattcactag aaaaggagca aaaggtttat gaaggggctc acagagcggt tgaaattgaa 540
gctctaacac cacactccag ctactgtgta gtggctgaaa tatatcagcc catgttagac 600
agaagaagtc agagaagtga agagagatgt gtggaaattc catga 645

<210> 19
<211> 696
<212> DNA

<213> Homo sapiens

<400> 19
atgatgccta aacattgctt tctaggcttc ctcatcagtt tcttccttac tggtgtagca 60
ggaactcagt caacgcatga gtctctgaag cctcagaggg tacaatttca gtcccgaaat 120
tttcacaaca ttttgcaatg gcagcctggg agggcactta ctggcaacag cagtgtctat 180
tttgtgcagt acaaaatata tggacagaga caatggaaaa ataaagaaga ctgttggggt 240
actcaagaac tctcttgtga ccttaccagt gaaacctcag acatacagga accttattac 300
gggagggtga gggcggcctc ggctgggagc tactcagaat ggagcatgac gccgcggttc 360
actccctggt gggaaacaaa aatagatcct ccagtcatga atataaccca agtcaatggc 420
tctttgttgg taattctcca tgctccaaat ttaccatata gataccaaaa ggaaaaaaat 480
gtatctatag aagattacta tgaactacta taccgagttt ttataattaa caattcacta 540
gaaaaggagc aaaaggttta tgaaggggct cacagagcgg ttgaaattga agctctaaca 600
ccacactcca gctactgtgt agtggctgaa atatatcagc ccatgttaga cagaagaagt 660
cagagaagtg aagagagatg tgtggaaatt ccatga 696

<210> 20
<211> 792
<212> DNA
<213> Homo sapiens

<400> 20
atgatgccta aacattgctt tctaggcttc ctcatcagtt tcttccttac tggtgtagca 60
ggaactcagt caacgcatga gtctctgaag cctcagaggg tacaatttca gtcccgaaat 120
tttcacaaca ttttgcaatg gcagcctggg agggcactta ctggcaacag cagtgtctat 180
tttgtgcagt acaaaatcat gttctcatgc agcatgaaaa gctctcacca gaagccaagt 240
ggatgctggc agcacatttc ttgtaacttc ccaggctgca gaacattggc taaatatgga 300
cagagacaat ggaaaaataa agaagactgt tggggtactc aagaactctc ttgtgacctt 360
accagtgaaa cctcagacat acaggaacct tattacggga gggtgagggc ggcctcggct 420
gggagctact cagaatggag catgacgccg cggttcactc cctggtggga aacaaaaata 480
gatcctccag tcatgaatat aacccaagtc aatggctctt tgttggtaat tctccatgct 540
ccaaatttac catatagata ccaaaaggaa aaaaatgtat ctatagaaga ttactatgaa 600
ctactatacc gagtttttat aattaacaat tcactagaaa aggagcaaaa ggtttatgaa 660
ggggctcaca gagcggttga aattgaagct ctaacaccac actccagcta ctgtgtagtg 720
gctgaaatat atcagcccat gttagacaga agaagtcaga gaagtgaaga gagatgtgtg 780
gaaattccat ga 792

<210> 21
<211> 780
<212> DNA
<213> Homo sapiens

<400> 21
atgtatgtat tatctccagt ggaatttata attctacaac ttttatttat tcaggccatt 60
tccagcagtt taaaaggttt cctttcagct atgagactgg ctcatagagg ctgtaatgtt 120
gatacaccag tttcaacgct cacaccagtg aagacttcag aatttgaaaa ctttaaaact 180
aaaatggtta tcacatccaa aaaagactat cctctaagta agaattttcc atattccttg 240
gaacatcttc agacttctta ctgtgggctt gtccgagttg atatgcgtat gctttgctta 300
aaaagcctta ggaaattaga cttgagtcac aaccatataa aaaagcttcc agctacaatt 360
ggagacctca tacaccttca agaacttaac ctgaatgaca atcacttgga gtcatttagt 420
gtagccttgt gtcattctac actccagaag tcacttcgga gtttggacct cagcaagaac 480
aaaatcaagg cactccctgt gcagttttgc cagctccagg aacttaagaa tttaaaactt 540
gacgataatg aattgattca atttccttgc aagataggac aactaataaa ccttcgcttt 600
ttgtcagcag ctcgaaataa gcttccattt ttgcctagtg aatttagaaa tttatccctt 660
gaatacttgg atctttttgg aaatactttt gaacaaccaa aagtccttcc agtaataaag 720
ctgcaagcac cattaacttt attggaatct tctgcacgaa ccatattaca taataggtaa 780

<210> 22

```
Asn Pro Phe Pro Lys Pro Gly Glu Ser Glu Ile Phe Ile Asp Ser Ile
            180                 185                 190
Arg Arg Leu Glu Asn Leu Tyr Val Val Glu Glu Lys Asp Leu Cys Ala
        195                 200                 205
Ala Cys Leu Arg Lys Cys Gln Asn Ala Arg Asp Asn Leu Asn Arg Ile
    210                 215                 220
Lys Asn Met Ala Thr Thr Thr Pro Arg Lys Thr Ile Phe Pro Asn Leu
225                 230                 235                 240
Ile Ser Pro Asn Ser Met Ala Lys Asp Ser Trp Glu Asp Trp Arg
                245                 250                 255


<210> 40
<211> 214
<212> PRT
<213> Homo sapiens

<400> 40
Met Gln Ala Gly Thr Gln Ser Thr His Glu Ser Leu Lys Pro Gln Arg
 1               5                  10                  15
Val Gln Phe Gln Ser Arg Asn Phe His Asn Ile Leu Gln Trp Gln Pro
            20                  25                  30
Gly Arg Ala Leu Thr Gly Asn Ser Ser Val Tyr Phe Val Gln Tyr Lys
        35                  40                  45
Ile Tyr Gly Gln Arg Gln Trp Lys Asn Lys Glu Asp Cys Trp Gly Thr
    50                  55                  60
Gln Glu Leu Ser Cys Asp Leu Thr Ser Glu Thr Ser Asp Ile Gln Glu
65                  70                  75                  80
Pro Tyr Tyr Gly Arg Val Arg Ala Ala Ser Ala Gly Ser Tyr Ser Glu
                85                  90                  95
Trp Ser Met Thr Pro Arg Phe Thr Pro Trp Trp Glu Thr Lys Ile Asp
            100                 105                 110
Pro Pro Val Met Asn Ile Thr Gln Val Asn Gly Ser Leu Leu Val Ile
        115                 120                 125
Leu His Ala Pro Asn Leu Pro Tyr Arg Tyr Gln Lys Glu Lys Asn Val
    130                 135                 140
Ser Ile Glu Asp Tyr Tyr Glu Leu Leu Tyr Arg Val Phe Ile Ile Asn
145                 150                 155                 160
Asn Ser Leu Glu Lys Glu Gln Lys Val Tyr Glu Gly Ala His Arg Ala
                165                 170                 175
Val Glu Ile Glu Ala Leu Thr Pro His Ser Ser Tyr Cys Val Val Ala
            180                 185                 190
Glu Ile Tyr Gln Pro Met Leu Asp Arg Arg Ser Gln Arg Ser Glu Glu
        195                 200                 205
Arg Cys Val Glu Ile Pro
    210


<210> 41
<211> 231
<212> PRT
<213> Homo sapiens

<400> 41
Met Met Pro Lys His Cys Phe Leu Gly Phe Leu Ile Ser Phe Phe Leu
 1               5                  10                  15
Thr Gly Val Ala Gly Thr Gln Ser Thr His Glu Ser Leu Lys Pro Gln
            20                  25                  30
Arg Val Gln Phe Gln Ser Arg Asn Phe His Asn Ile Leu Gln Trp Gln
```

```
        35                  40                  45
Pro Gly Arg Ala Leu Thr Gly Asn Ser Ser Val Tyr Phe Val Gln Tyr
    50                  55                  60
Lys Ile Tyr Gly Gln Arg Gln Trp Lys Asn Lys Glu Asp Cys Trp Gly
65                  70                  75                  80
Thr Gln Glu Leu Ser Cys Asp Leu Thr Ser Glu Thr Ser Asp Ile Gln
                85                  90                  95
Glu Pro Tyr Tyr Gly Arg Val Arg Ala Ala Ser Ala Gly Ser Tyr Ser
            100                 105                 110
Glu Trp Ser Met Thr Pro Arg Phe Thr Pro Trp Trp Glu Thr Lys Ile
        115                 120                 125
Asp Pro Pro Val Met Asn Ile Thr Gln Val Asn Gly Ser Leu Leu Val
    130                 135                 140
Ile Leu His Ala Pro Asn Leu Pro Tyr Arg Tyr Gln Lys Glu Lys Asn
145                 150                 155                 160
Val Ser Ile Glu Asp Tyr Tyr Glu Leu Leu Tyr Arg Val Phe Ile Ile
                165                 170                 175
Asn Asn Ser Leu Glu Lys Glu Gln Lys Val Tyr Glu Gly Ala His Arg
            180                 185                 190
Ala Val Glu Ile Glu Ala Leu Thr Pro His Ser Ser Tyr Cys Val Val
        195                 200                 205
Ala Glu Ile Tyr Gln Pro Met Leu Asp Arg Arg Ser Gln Arg Ser Glu
    210                 215                 220
Glu Arg Cys Val Glu Ile Pro
225                 230
```

<210> 42
<211> 263
<212> PRT
<213> Homo sapiens

<400> 42

```
Met Met Pro Lys His Cys Phe Leu Gly Phe Leu Ile Ser Phe Phe Leu
 1               5                  10                  15
Thr Gly Val Ala Gly Thr Gln Ser Thr His Glu Ser Leu Lys Pro Gln
            20                  25                  30
Arg Val Gln Phe Gln Ser Arg Asn Phe His Asn Ile Leu Gln Trp Gln
        35                  40                  45
Pro Gly Arg Ala Leu Thr Gly Asn Ser Ser Val Tyr Phe Val Gln Tyr
    50                  55                  60
Lys Ile Met Phe Ser Cys Ser Met Lys Ser Ser His Gln Lys Pro Ser
65                  70                  75                  80
Gly Cys Trp Gln His Ile Ser Cys Asn Phe Pro Gly Cys Arg Thr Leu
                85                  90                  95
Ala Lys Tyr Gly Gln Arg Gln Trp Lys Asn Lys Glu Asp Cys Trp Gly
            100                 105                 110
Thr Gln Glu Leu Ser Cys Asp Leu Thr Ser Glu Thr Ser Asp Ile Gln
        115                 120                 125
Glu Pro Tyr Tyr Gly Arg Val Arg Ala Ala Ser Ala Gly Ser Tyr Ser
    130                 135                 140
Glu Trp Ser Met Thr Pro Arg Phe Thr Pro Trp Trp Glu Thr Lys Ile
145                 150                 155                 160
Asp Pro Pro Val Met Asn Ile Thr Gln Val Asn Gly Ser Leu Leu Val
                165                 170                 175
Ile Leu His Ala Pro Asn Leu Pro Tyr Arg Tyr Gln Lys Glu Lys Asn
            180                 185                 190
Val Ser Ile Glu Asp Tyr Tyr Glu Leu Leu Tyr Arg Val Phe Ile Ile
        195                 200                 205
```

```
Asn Asn Ser Leu Glu Lys Glu Gln Lys Val Tyr Glu Gly Ala His Arg
    210                 215                 220
Ala Val Glu Ile Glu Ala Leu Thr Pro His Ser Ser Tyr Cys Val Val
225                 230                 235                 240
Ala Glu Ile Tyr Gln Pro Met Leu Asp Arg Arg Ser Gln Arg Ser Glu
                245                 250                 255
Glu Arg Cys Val Glu Ile Pro
            260
```

<210> 43
<211> 259
<212> PRT
<213> Homo sapiens

<400> 43

```
Met Tyr Val Leu Ser Pro Val Glu Phe Ile Ile Leu Gln Leu Leu Phe
 1               5                  10                  15
Ile Gln Ala Ile Ser Ser Ser Leu Lys Gly Phe Leu Ser Ala Met Arg
            20                  25                  30
Leu Ala His Arg Gly Cys Asn Val Asp Thr Pro Val Ser Thr Leu Thr
        35                  40                  45
Pro Val Lys Thr Ser Glu Phe Glu Asn Phe Lys Thr Lys Met Val Ile
    50                  55                  60
Thr Ser Lys Lys Asp Tyr Pro Leu Ser Lys Asn Phe Pro Tyr Ser Leu
65                  70                  75                  80
Glu His Leu Gln Thr Ser Tyr Cys Gly Leu Val Arg Val Asp Met Arg
                85                  90                  95
Met Leu Cys Leu Lys Ser Leu Arg Lys Leu Asp Leu Ser His Asn His
            100                 105                 110
Ile Lys Lys Leu Pro Ala Thr Ile Gly Asp Leu Ile His Leu Gln Glu
        115                 120                 125
Leu Asn Leu Asn Asp Asn His Leu Glu Ser Phe Ser Val Ala Leu Cys
    130                 135                 140
His Ser Thr Leu Gln Lys Ser Leu Arg Ser Leu Asp Leu Ser Lys Asn
145                 150                 155                 160
Lys Ile Lys Ala Leu Pro Val Gln Phe Cys Gln Leu Gln Glu Leu Lys
                165                 170                 175
Asn Leu Lys Leu Asp Asp Asn Glu Leu Ile Gln Phe Pro Cys Lys Ile
            180                 185                 190
Gly Gln Leu Ile Asn Leu Arg Phe Leu Ser Ala Ala Arg Asn Lys Leu
        195                 200                 205
Pro Phe Leu Pro Ser Glu Phe Arg Asn Leu Ser Leu Glu Tyr Leu Asp
    210                 215                 220
Leu Phe Gly Asn Thr Phe Glu Gln Pro Lys Val Leu Pro Val Ile Lys
225                 230                 235                 240
Leu Gln Ala Pro Leu Thr Leu Leu Glu Ser Ser Ala Arg Thr Ile Leu
                245                 250                 255
His Asn Arg
```

<210> 44
<211> 416
<212> PRT
<213> Homo sapiens

<400> 44

```
Met Lys Leu His Cys Glu Val Glu Val Ile Ser Arg His Leu Pro Ala
```